Original Article

How Happy Is Happy Enough?

A Cross-Cultural Comparison of Optimal Cut Points for the Positive Mental Health Scale

Gabriel Bonnin[1,2], Gerrit Hirschfeld[3], Ruth von Brachel[1,2], and Jürgen Margraf[1,2]

[1]Mental Health Research and Treatment Center, Ruhr University Bochum, Germany
[2]German Center for Mental Health (DZPG), Bochum, Germany
[3]Faculty of Business, University of Applied Sciences Bielefeld, Germany

Abstract: As positive mental health (PMH) has a significant impact on general and mental health, it is an important target for interventions. Cut points are a useful basis for identifying participants with a greater need for such interventions. Representative (n = 9,440) and student (n = 22,833) samples from Germany, Russia, the US, and China were reanalyzed. Two different anchors were used to determine optimal cut points for the PMH-scale: (1) a combined measure of PMH-related questionnaires and (2) the Global Assessment of Functioning Scale (GAF). A kernel-based method to determine optimal cut points and bootstrapping to identify potential cross-cultural differences were used. Acceptable to excellent levels of classification accuracy were found regarding the combined measure (AUCs between .75 and .87) across all samples. Using the GAF resulted in poor discriminatory power (AUC = .69). Optimal cut points varied systematically between countries and samples. Country and sample-specific cut points for the PMH scale should be used to identify individuals with high versus low levels of PMH. Specifically, we suggest using cut points of 21, 22, and 24 in Germany, Russia, and the US, respectively. For student samples, we recommend cut points of 18, 19, and 20 in Germany, Russia, and China, respectively.

Keywords: positive mental health, optimal cut points, bootstrapping

In the past, mental health care and research primarily focused on reducing mental disorders. However, there is now a growing emphasis on promoting positive mental health (PMH), which encompasses emotional, psychological, and social well-being (Keyes, 2005). Studies indicate that PMH and related constructs significantly impact the development and course of mental disorders beyond mere psychopathological indicators (Tennant, Goens, et al., 2007), making PMH a crucial target for intervention and prevention. While interventions to promote PMH are well established (Bolier et al., 2013), identifying individuals who would most benefit from these interventions remains a challenge. This study aims to determine optimal cut points for a measure of PMH.

Several instruments assess different facets of PMH, including the Warwick-Edinburgh Mental Well-Being Scale (Tennant, Hiller, et al., 2007), the Mental Health Continuum (Keyes et al., 2002), the Euthymia Scale (Fava & Bech, 2016), and the Positive Mental Health Scale (Lukat et al., 2016). The Positive Mental Health Scale (PMH-scale) is a short, unidimensional measure primarily capturing the psychological and emotional components of PMH in a general, cross-situational, and person-centered way (Lukat et al., 2016). While this unidimensional perspective may not fully encompass the breadth of factors influencing mental health, its brevity significantly reduces the measurement burden, making it highly effective for epidemiological and interventional research where efficiency and ease of administration are crucial.

Most measures of PMH provide continuous scores, requiring cut points to interpret individual results or make informed decisions. Cut points have numerous potential applications in PMH. They help identify individuals who may benefit from interventions, facilitate the monitoring and interpretation of PMH across populations, and enable health professionals and researchers to gauge mental health trends and patterns. Additionally, PMH is a protective factor against suicidality (Brailovskaia et al., 2019). Thus, assessing and monitoring PMH is essential in clinical settings to evaluate the risk of suicidality and determine the necessity for interventions. While there are many measures of psychopathology with well-established cut points (von Glischinski et al., 2019), there are – to our knowledge – no such cut points for assessing of PMH.

Therefore, the aim of this study is twofold: first, we want to determine optimal cut points for the PMH-scale based on representative and student samples by using self- and clinician-rated measures as anchors for classification. Second, we want to compare optimal cut points cross-culturally between German, US, Russian, and Chinese participants. While the PMH-scale has been shown to measure a conceptually equivalent construct in different cultures (Bieda et al., 2017; Velten et al., 2021), it is still unclear whether the optimal cut points for the PMH-scale vary between these countries.

Methods

Participants

We reanalyzed a subset of data derived from the Bochum Optimism and Mental Health Study (BOOM; Margraf, Zhang, et al., 2020) and the Dresden Predictor Study (DPS; Trumpf et al., 2010).

The BOOM project investigates protective and risk factors for mental health across different countries using a combination of epidemiological, cross-sectional, and longitudinal study designs (Margraf, Zhang, et al., 2020). The study received ethical approval from the local ethics committee of the faculty of psychology at the Ruhr University Bochum (approval number: 20110512).

The DPS was a prospective epidemiological study of mental disorders and received ethical approval from the Office for Data Protection in Saxony and the State of Saxony Public Health Association. For more details on the study, see Trumpf and colleagues (2010).

BOOM Samples

To estimate and compare optimal cut points in different populations we used representative (N = 8,065) and student (N = 22,833) samples from the BOOM study.

Representative Samples

The representative samples were collected in Germany (n = 1,877), the USA (n = 2,791), and Russia (n = 2,634). Systematic sampling ensured the representativeness of adult residents using 2011 census data, considering age, gender, and education. Data were collected between November 2012 and February 2014 using telephone interviews. Participants were informed about the study's purpose and assured of their anonymity before providing written informed consent to participate.

Student Samples

Student samples were gathered in Germany (n = 4,575), China (n = 13,163), and Russia (n = 3,905). In Germany, students were recruited from the Ruhr University Bochum in 2011 through an online questionnaire. Chinese participants were recruited from Capital Normal University Beijing, Hebei United University, and Nanjing University in 2012. Russian students were recruited from Lomonosov University Moscow, Voronezh State University, and Orenburg State Medical University in 2012. Data collection in China and Russia was conducted in group testing sessions using either online or paper-pencil questionnaires.

DPS Sample

The DPS sample included a random selection of young women aged 18–25 from the 1996 population registers of Dresden, Germany (Trumpf et al., 2010). A baseline survey was conducted from July 1996 to September 1997, followed by a subsequent assessment 17 months later. At both times, the German version of the Anxiety Disorder Interview Schedule-Lifetime (ADIS-IV-L; Margraf et al., 1996) was administered, and several self-report questionnaires were completed. The present study included data from participants who took part in both assessments (N = 1,375).

Measures

Positive Mental Health Scale (PMH-Scale)

The PMH-scale (Lukat et al., 2016) measures emotional, psychological, and social aspects of well-being through nine self-report statements (e.g., "I enjoy my life") on a 4-point Likert scale ranging from 0 (= *disagree*) to 3 (= *agree*). The sum score of the item responses (range = 0–27) indicates higher levels of positive mental health. The scale has been confirmed as a unidimensional instrument with high internal consistency (Cronbach's α = .93), good retest-reliability (Pearson correlation = .74–.81), and scalar invariance across samples and over time (Lukat et al., 2016). Additionally, it has demonstrated good convergent and discriminant validity and sensitivity to therapeutic change in diverse samples (Lukat et al., 2016).

Satisfaction With Life Scale (SWLS)

The SWLS (Glaesmer et al., 2011) measures satisfaction with one's life using five items (e.g., "In most ways my life corresponds to my ideal") on a 7-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*). The total score (range = 5–35) is the sum of the item responses, with higher scores indicating greater life satisfaction. The SWLS has good psychometric properties, showing convergent and discriminant validity across samples (Pavot & Diener, 2008) and cross-cultural measurement invariance (Tucker et al., 2006).

Sense of Coherence Scale (SOC-L9)

Sense of coherence was measured with a short version of the sense of coherence scale (Schumacher et al., 2000).

The SOC-L9 consists of nine items (e.g., "Do you have the feeling that you are in an unfamiliar situation and don't know what to do?") rated on a 7-point Likert scale, with varying anchor wording depending on the item (e.g., 1 = *very often*, 7 = *very rare or never*). After recoding inverse-poled items, the item responses are summed up (range = 9–63), with higher scores indicating a greater sense of coherence. The SOC-L9 captures the sense of coherence in all its aspects and has good internal consistency (Cronbach's $\alpha = .87$).

Resilience Scale (RS-11)

Stress resilience was assessed using the short version of the Wagnild and Young Resilience Scale (Schumacher et al., 2005). The RS-11 consists of 11 statements (e.g., "I usually manage one way or another"), which are rated on a scale from 1 (= *I disagree*) to 7 (= *I agree*). Higher item sum scores (range = 11–77) indicate greater levels of resilience. The RS-11 showed good reliability and convergent validity in a German sample (Schumacher et al., 2005).

Subjective Happiness Scale (SHS)

The SHS (Lyubomirsky & Lepper, 1999) measures global subjective happiness using four items (e.g., "In general, I consider myself:...") on a 7-point Likert scale, with different verbal anchors depending on the item (e.g., 1 = *not happy*, 7 = *very happy*). The responses are summed up (range = 4–28) with higher total scores indicating greater subjective happiness. The SHS has shown good internal consistency and convergent and discriminant validity in several countries (Swami et al., 2009).

Questionnaire-Social Support (F-SozU K-14)

The F-SozU K-14 (Fydrich et al., 2009) measures perceived social support using 14 items (e.g., "I experience a lot of understanding and security from others") on a 5-point Likert scale ranging from 1 (= *not true*) to 5 (= *true*). Item responses are summed up (range = 14–70), with higher scores indicating greater perceived social support. The F-SozU K-14 showed excellent internal consistency and good convergent and discriminant validity in a German sample (Fydrich et al., 2009).

Global Assessment of Functioning (GAF)

The GAF (Hall, 1995) is a clinician-rated measure of overall psychological, social, and occupational functioning, as well as well-being, covering the range from severe psychopathology (1) to positive mental health (100). The 100-point scale consists of 10 intervals, with verbal anchors for each interval to facilitate the interviewer's assessment of the patient's level of functioning. In our study, the GAF was rated by the clinician at the end of a structured clinical interview, which lasted between 30 and 90 minutes, depending on the complexity of the participants' complaints.

Data Analysis

Means and standard deviations of demographic variables, reference measures (SWLS, SOC-L9, RS-11, SHS, F-SozU K-14, GAF), and PMH scores were calculated separately for each of the samples.

We used two different binary criteria as classification anchors:

First, we used an integrated measure based on Keyes' definition of PMH (Keyes et al., 2002), encompassing related constructs such as a sense of coherence, resilience, happiness, and social support. Participants who scored above the scale midpoint in all individual scales were grouped into the positive group (SWLS>20 & SOC-9>28 & RS-11>44 & SHS>16 & F-SozU K-14>42), while those who scored at or below the midpoint were grouped into the negative group (SWLS$\leq$20 & SOC-9$\leq$28 & RS-11$\leq$44 & SHS$\leq$16 & F-SozU K-14$\leq$42). For the SWLS, using the scale midpoint for classification is supported by literature (Pavot & Diener, 1993). In the absence of established cut points for the other scales, we applied the same rationale for categorizing participants across all self-report measures.

Second, we used the GAF (Hall, 1995) as a clinician-rated measure of functioning and well-being. This approach minimizes the risk of overestimating classification accuracy, which can occur if both the anchor and the instrument for which optimal cut points are to be determined are self-report measures (Hirschfeld et al., 2020). A cut-off score (GAF>80) was chosen for dichotomization, as the verbal anchors of the GAF scores greater than 80 reflect general satisfaction with life ("Absent or minimal symptoms (e.g., mild anxiety before an exam), good functioning in all areas, interested and involved in a wide range of activities, socially effective, generally satisfied with life, no more than everyday problems or concerns (e.g., an occasional argument with family members").

For each of the two anchors optimal cut points (OC) were estimated based on maximizing the Youden's index, that is, finding the cut points of the PMH-scale that represent a trade-off between correctly identifying true positive and avoiding false positive cases. To determine the variability of the OC and to avoid overestimating classification accuracy, we applied a bootstrap resampling procedure with 1,000 repetitions. Furthermore, confidence intervals (CI) for OC and out-of-bag estimates for measures of classification accuracy (area under the receiver operating characteristic curve, sensitivity, specificity) were calculated. The area under the receiver operating characteristic curve (AUC) was interpreted using guidelines provided by Hosmer and colleagues (Hosmer et al., 2013).

Table 1. Descriptive statistics

	Representative samples (BOOM)			Student samples (BOOM)			DPS sample
	Germany	USA	Russia	Germany	China	Russia	Germany
N	1,877	2,791	2,634	4,575	13,163	3,905	1,375
Age; *M* (*SD*)	51.4 (17.3)	54.4 (17.4)	41.9 (16.7)	26.5 (4.01)	19.7 (1.85)	19.8 (2.37)	22.7 (1.80)
Percent female	59.0%	58.4%	53.4%	55.0%	61.7%	64.5%	100%
PMH-scale; *M* (*SD*)	21.8 (4.82)	23.2 (5.09)	20.8 (5.32)	18.0 (6.00)	21.1 (5.07)	19.0 (5.14)	20.6 (4.49)
SWLS; *M* (*SD*)	27.0 (5.85)	27.1 (6.73)	23.2 (6.79)	25.1 (6.57)	23.9 (6.48)	24.3 (5.80)	–
SOC-9; *M* (*SD*)	49.8 (8.25)	52.0 (9.67)	49.3 (9.12)	42.6 (9.84)	45.1 (7.56)	43.5 (8.02)	–
RS-11; *M* (*SD*)	64.8 (9.14)	66.4 (11.5)	65.1 (10.3)	55.6 (14.0)	58.4 (8.49)	59.2 (8.84)	–
SHS; *M* (*SD*)	21.6 (4.27)	22.1 (4.72)	19.8 (5.04)	18.9 (5.41)	21.7 (4.30)	20.4 (4.42)	–
F-SozU K-14	63.4 (8.30)	59.0 (12.1)	61.0 (9.75)	59.5 (11.0)	56.6 (12.1)	57.9 (11.4)	–
% positive (combined scales)	78.5%	72.7%	52.8%	54.6%	60.0%	61.7%	–
GAF; *M* (*SD*)	–	–	–	–	–	–	86.6 (10.0)
% GAF > 80	–	–	–	–	–	–	70.1%

Note. PMH-scale = Positive mental health scale, SWLS=Satisfaction with life scale, SOC-9 = Sense of coherence scale, RS-11 = Resilience scale, SHS = Subjective happiness scale, F-SozU K-14 = Questionnaire-Social Support, GAF = Global assessment of functioning scale.

All analyses were run in R (R Core Team, 2022) using the cut points package (Thiele & Hirschfeld, 2021). The anonymized dataset (Bonnin et al., 2024b) and R code (Bonnin et al., 2024a) are available online.

Analyses regarding reliability, factorial validity, measurement invariance, and the relationship between self-report scales used in this study, which are beyond the scope of this article, can be found in the Electronic Supplementary Materials, ESM 1.

Results

Descriptive Statistics

In the representative BOOM samples, PMH was highest in the US (M = 23.2), followed by Germany (M = 21.8) and Russia (M = 20.8). In the student samples, PMH was highest among Chinese students (M = 21.1), followed by Russian (M = 19.0) and German students (M = 18.0). In the DPS sample (M = 20.6), PMH was comparable to the German representative BOOM sample. Further descriptive statistics of demographic variables, PMH scores, and reference measures for all samples can be found in Table 1.

Optimal Cut Points and Classification Accuracy

Optimal cut points and measures of classification accuracy for all three anchors are summarized in Table 2.

Classification Accuracy Against the Combined Scales

In the representative samples, using the combined scales as an anchor for classification, the OC for the PMH-scale differed between cultures. In the German sample, the OC was 21 (95% CI 19–22), while participants from the USA and Russia had higher OC of 24 (95% CI 23–24) and 22 (95% CI 19–23), respectively. The highest classification accuracy was obtained in the German sample with an AUC of .82, indicating excellent discriminatory power. For the USA (AUC = .77) and Russia (AUC = .75), acceptable levels of classification accuracy were found. Classification sensitivity was good in the German sample, acceptable in the US sample, and poor in the Russian sample. Specificity was poor across Germany, the US, and Russia (Table 2; Figure 1).

A similar pattern emerged for the student samples using classification based on the combination of scales. Optimal cut points were 18 (95% CI 17–18) for German, 20 (95% CI 20–21) for Chinese, and 19 (95% CI 18–19) for Russian students. Classification accuracy was excellent in German (AUC = .87) and Russian (AUC = .81) student samples, and acceptable among Chinese students (AUC = .76). Sensitivity was good among German and Chinese, and acceptable among Russian students. Specificity was acceptable in German and Russian while being poor in the Chinese student sample (Table 2; Figure 2).

Classification Accuracy Against the GAF

The use of the clinician-rated GAF scale as an anchor resulted in poor discriminatory power (AUC = .69), as well as poor sensitivity and specificity estimates. The cut point identified as optimal was 21 (95% CI 19–24) (Table 2; Figure 3).

Discussion

The present study aimed to determine optimal cut points for the PMH-scale using self-report and clinician-rated

Table 2. Optimal cut points and classification accuracy

	Representative samples (BOOM)			Student samples (BOOM)			DPS sample
Variable	Germany	USA	Russia	Germany	China	Russia	Germany
Combined scales							–
AUC	.82	.77	.75	.87	.76	.81	–
OC [95% CI]	21 [19;22]	24 [23;24]	22 [19;23]	18 [17;18]	20 [20;21]	19 [18;19]	–
% high PMH	67.9%	64.4%	51.4%	51.2%	63.6%	56.8%	
Sensitivity [oob]	.82	.77	.69	.85	.80	.76	–
Specificity [oob]	.67	.68	.66	.76	.61	.72	–
GAF (> 80)							
AUC	–	–	–	–	–	–	.69
OC [95% CI]	–	–	–	–	–	–	21 [19;24]
% high PMH							56.7%
Sensitivity [oob]	–	–	–	–	–	–	.64
Specificity [oob]	–	–	–	–	–	–	.62

Note. AUC = area under the curve; OC = optimal cut point; CI = confidence interval; oob = out-of-bag estimate; combined scales = SWLS, SOC-L9, RS-11, SHS; F-SozU K-14; GAF = Global assessment of functioning scale.

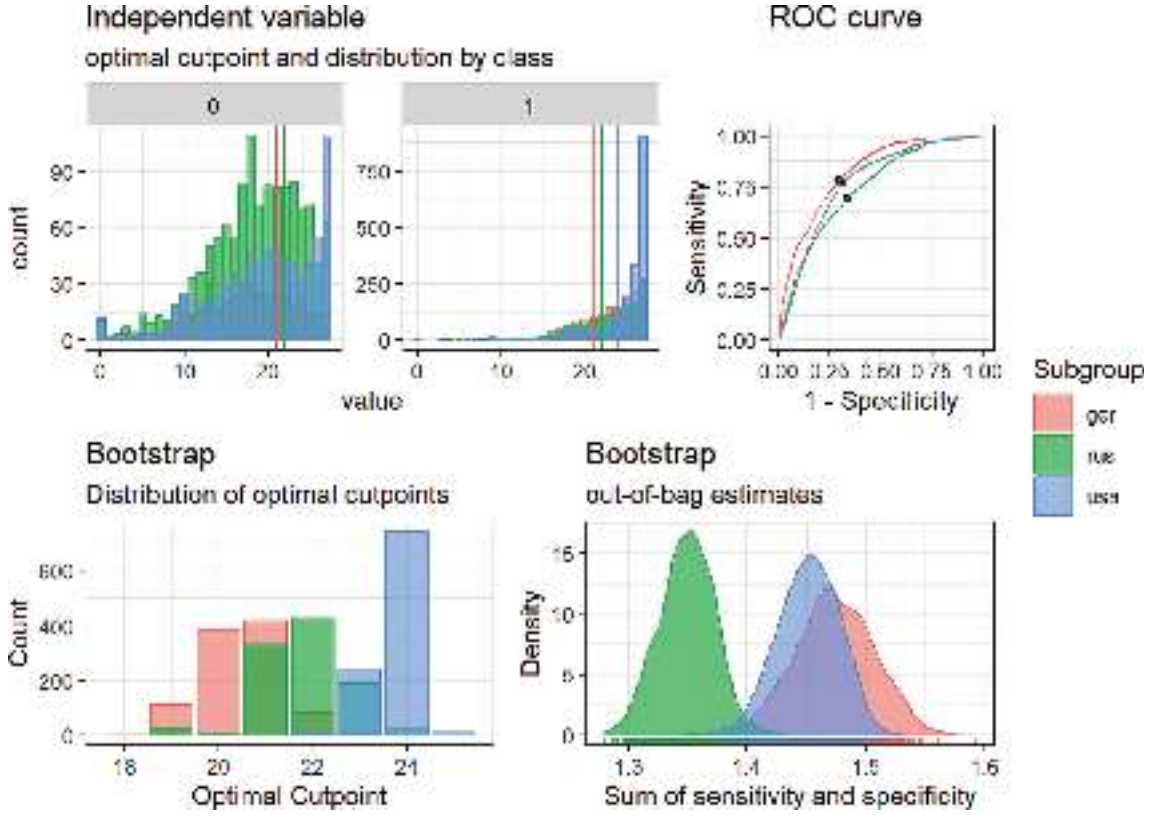

Figure 1. Optimal cut points for the PMH-scale regarding the combined scales (SWLS, SOC-L9, RS-11, SHS, F-SozU K-14) in three representative BOOM samples (Germany, Russia, USA).

measures as anchors, and to compare these cut points across representative and student samples from Germany, the USA, Russia, and China.

We will first discuss the OC for the different classification anchors, and then examine the differences in OC between countries and sample types. Finally, we address the limitations and strengths of the current study and provide an outlook on areas for future research.

Using a combination of measures associated with PMH, receiver operating characteristic curve analysis of the

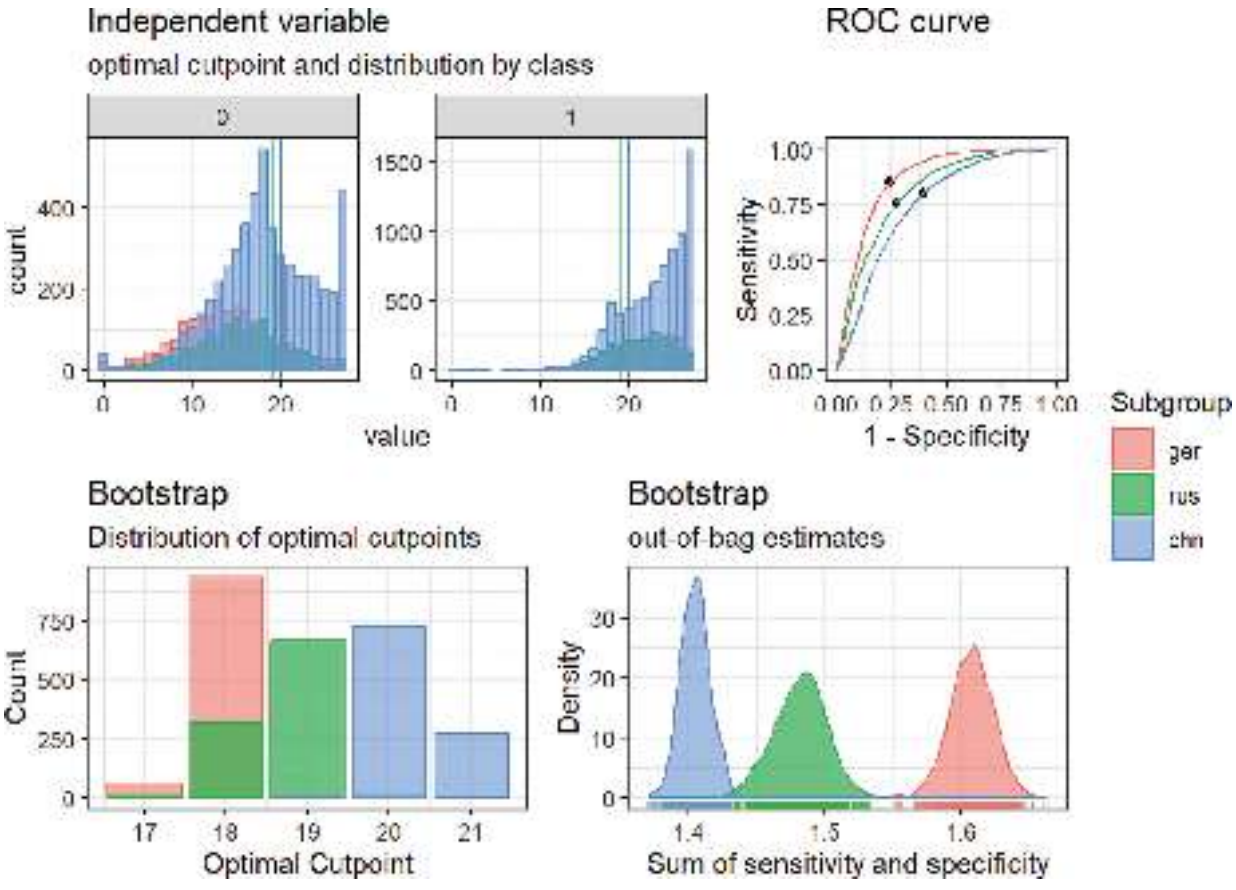

Figure 2. Optimal cut points for the PMH-scale regarding the combined scales (SWLS, SOC-L9, RS-11, SHS, F-SozU K-14) in three student BOOM samples (Germany, Russia, China).

PMH-scale revealed an acceptable to excellent ability to discriminate between participants with low and high scores. Optimal cut points ranged from 21 (Germany) to 24 (USA) in the representative samples, and from 18 (Germany) to 20 (China) in the student samples.

Regarding the clinician-rated GAF, the OC identified was 21, consistent with the results of the German representative sample. However, the PMH-scale did not sufficiently discriminate between high and low levels of functioning and well-being. This result could be due to several reasons: First, it is possible that the classification accuracy of the self-report-based anchors was overestimated in comparison to clinician ratings due to the differing modes of administration of the PMH-scale and the GAF scale. The PMH-scale employs a self-report approach, whereas the GAF scale utilizes a clinical interview. Second, psychometric issues concerning the reliability (Hilsenroth et al., 2000; Söderberg et al., 2005), as well as the concurrent (Hilsenroth et al., 2000) and predictive validity (Bacon et al., 2002) of the GAF, might have negatively impacted the classification accuracy in our study. Third, the GAF is a unidimensional measure of psychopathology and well-being. This view of mental health has prevailed in clinical practice and research for decades, but it is increasingly contested by more complex multidimensional models. For example, the dual-factor model of mental health distinguishes between mental health and disorder as two correlated yet distinct dimensions (Suldo & Shaffer, 2008). Furthermore, numerous studies have indicated that individuals can concurrently exhibit high levels of PMH and psychopathology (Brailovskaia et al., 2022; Teismann et al., 2018). Given that the PMH-scale is solely focused on positive mental health, employing a unidimensional measure that also encompasses psychopathology may not be a suitable anchor for classification, resulting in a low discriminatory power. Taken together these results indicate that the self-report PMH-scale may not be able to substitute clinicians' ratings on an individual level.

Secondly, we investigated cross-cultural differences. Optimal cut points for the PMH-scale showed systematic differences across countries, beyond random fluctuations within the samples. Optimal cut points were generally lowest in Germany, medium in Russia, and highest in China and the USA. Notably, the largest differences were observed between the German and US samples, both of which can be considered culturally Western. While the confidence intervals of the optimal cut points in the representative samples

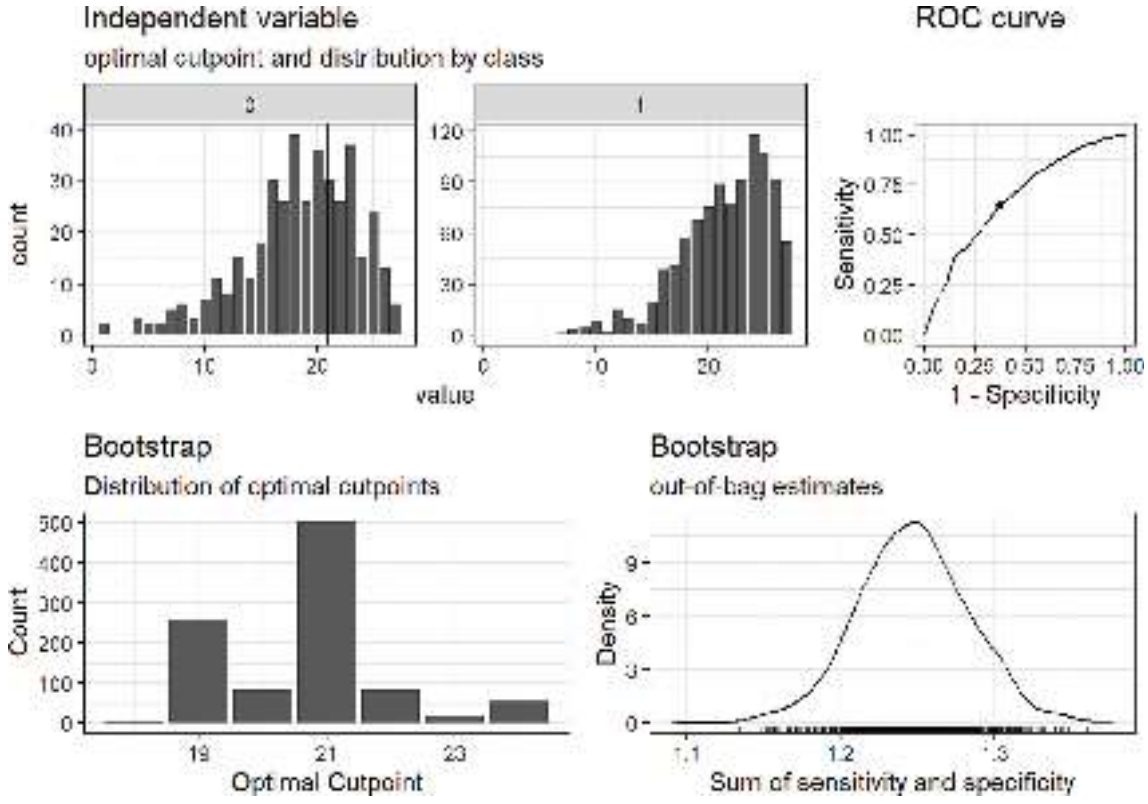

Figure 3. Optimal cut points for the PMH-scale regarding the GAF scale in the DPS sample.

of Russia and the USA as well as Russia and Germany showed some overlap, this was neither the case in the USA and Germany nor the student samples. Thus, we can conclude that the differences we observed are due to real differences in the way PMH is related to both anchors in these countries. In another study comparing PMH across countries during the COVID-19 pandemic, of the countries studied here, the highest PMH scores were found in the USA, followed by Germany and Russia (Margraf, Brailovskaia, et al., 2020), confirming the high levels in the USA found in our study. At the same time, the results of this study contradict our finding that PMH levels are higher in the Russian general population than in Germany. Given that PMH levels of the population can fluctuate over time and response to global crises, comparability with our results is limited.

Third, we investigated differences between student and representative samples. We found consistently lower cut points in student samples. Lower PMH scores in student samples compared to representative samples may indicate a positive association between PMH and age, supported by post hoc analyses in German and US samples. However, this association is variable across studies (Askari et al., 2021; Chuang et al., 2021; Lukat et al., 2016; Teismann & Brailovskaia, 2020). Additionally, cross-country comparisons reveal divergent trends in the age-PMH relationship. For example, PMH scores tend to decrease throughout the lifespan in Russia, whereas they exhibit a U-shaped trajectory in Germany (Schönfeld et al., 2017). Given these mixed findings, it is challenging to draw definitive conclusions about the implications of the discrepancies in optimal cut points between student and representative samples in our study. Furthermore, methodological variations might also account for differences between representative and student samples. Representative samples were assessed through telephone interviews, while student samples completed online questionnaires. Previous research has shown that the method of data collection can significantly influence responses to PMH measures (Zhang et al., 2017). Zhang and colleagues (2017) found substantial differences in PMH scores between online questionnaires and telephone interviews, with the latter typically resulting in more socially desirable responses (i.e., higher self-reported PMH). This indicates that mode effects may also contribute to the observed differences between the two sample groups, emphasizing the necessity for cautious interpretation of these results.

The present study has some limitations that should be noted: First, one of the two anchors was based on self-report, which may have led to an overestimation of classification accuracy. In particular, if both the anchor and the instrument for which optimal cut-off points are to

be determined are based on self-assessment, there is an increased risk that the classification accuracy will be overestimated (Hirschfeld et al., 2020). Second, none of the anchors used for classification relied on cut points already described in the scientific literature. Instead, we either used the scale midpoints or the verbal anchors for dichotomization. Although we provided a rationale for our decisions, they can still be considered somewhat arbitrary. Therefore, the validity of our results may be reduced. However, in our German representative samples, the self-rated and the clinician-rated anchors converged on an identical OC of 21 for the PMH-scale. Furthermore, the identified cut points are consistent with the results of large-scale epidemiological studies, such as the 2012 World Happiness Report (Helliwell et al., 2012). According to the report, 56% of Europe's population and 73% of the population in Northern America, Australia, and New Zealand fall within the 'thriving' category of the Cantril Ladder (Cantril, 1965). When applying our cut points to the representative BOOM data, 60% of the European (Germany and Russia) and 65% of the US sample are classified as having high levels of PMH. It is important to note that these percentages cannot be directly compared, as the data from the World Happiness Report is not available on the level of individual countries. However, they generally align with the classification according to our cut points. Third, the use of a unidimensional measure of psychopathology and well-being may have been an inappropriate choice, resulting in low discriminatory power for the clinician-rated GAF.

Besides these limitations, the study also has some notable strengths: First, we used large representative and student samples from different countries, which allowed the identification of cross-cultural differences in classification with the PMH-scale. Second, we employed different anchors based on various self-report measures and clinical interviews to determine classification accuracy. Third, the methodological approach was carefully chosen, accounting for random fluctuation of optimal cut points and avoiding estimation biases through bootstrapping.

The cut points found should be further validated, by incorporating more suitable and objective anchors for classification (e.g., asking participants whether they feel in need of an intervention). However, it remains a challenge to find a true "gold standard" criterion for the presence or absence of positive mental health. As states of mental health and disorder are socially constructed categories, the dichotomization of these dimensions remains to some extent artificial. Furthermore, the fit and utility of the cut points found should be investigated in future studies promoting PMH, that include individuals based on the optimal cut points identified here or at random.

Intervention decisions for individual patients should ideally be based on a multifactorial analysis. Apart from the methodological limitations discussed, the identified cut points should not be the only basis for these decisions. This study represents a first attempt to address a significant gap in assessment and clinical research by identifying individuals with low-level PMH, allowing progress toward more informed decision-making in this area.

Conclusion

We found that specific cut points for the PMH-scale should be used to identify individuals with high or low levels of PMH. Specifically, we suggest using cut points of 21, 22, and 24 in Germany, Russia, and the US, respectively. For student samples, we recommend cut points of 18, 19, and 20 in Germany, Russia, and China, respectively.

Electronic Supplementary Material

The following electronic supplementary material is available at https://doi.org/10.1027/1015-5759/a000850

ESM 1. ESM 1 contains results of suplementary analyses regarding reliability, factorial validity, measurement invariance, and the relationship between self-report scales used in this study.

References

Askari, F., Davoudi, I., Neysi, A., & Zargar, Y. (2021). Dream sharing and positive mental health in Iranian Culture. *Dreaming, 31*(3), 252–261. https://doi.org/10.1037/drm0000169

Bacon, S. F., Collins, M. J., & Plake, E. V. (2002). Does the Global Assessment of Functioning assess functioning? *Journal of Mental Health Counseling, 24*(3), 202–212.

Bieda, A., Hirschfeld, G., Schönfeld, P., Brailovskaia, J., Zhang, X. C., & Margraf, J. (2017). Universal happiness? Cross-cultural measurement invariance of scales assessing positive mental health. *Psychological Assessment, 29*(4), 408–421. https://doi.org/10.1037/pas0000353

Bolier, L., Haverman, M., Westerhof, G. J., Riper, H., Smit, F., & Bohlmeijer, E. (2013). Positive psychology interventions: A meta-analysis of randomized controlled studies. *BMC Public Health, 13*(1). https://doi.org/10.1186/1471–2458-13–119

Bonnin, G., Hirschfeld, G., von Brachel, R., & Margraf, J. (2024a). *How happy is happy enough? A cross-cultural comparison of optimal cut points for the Positive Mental Health Scale* [Code]. https://doi.org/10.23668/psycharchives.14646

Bonnin, G., Hirschfeld, G., von Brachel, R., & Margraf, J. (2024b). *How happy is happy enough? A cross-cultural comparison of optimal cut points for the Positive Mental Health Scale [Data]*. https://doi.org/10.23668/psycharchives.14647

Brailovskaia, J., Forkmann, T., Glaesmer, H., Paashaus, L., Rath, D., Schönfelder, A., Juckel, G., & Teismann, T. (2019). Positive mental health moderates the association between suicide ideation and suicide attempts. *Journal of Affective Disorders, 245*, 246–249. https://doi.org/10.1016/j.jad.2018.11.005

Brailovskaia, J., Teismann, T., Lewitzka, U., Gao, Z., Zhang, X. C., & Margraf, J. (2022). Suicidal ideation, suicide attempts and

positive mental health in Chinese medical students. *Journal of Affective Disorders Reports, 9*(March), 100354. https://doi.org/10.1016/j.jadr.2022.100354

Cantril, H. (1965). *The pattern of human concerns*. Rutgers University Press.

Chuang, S. P., Wu, J. Y. W., & Wang, C. S. (2021). Humor styles moderate the relationship between rumination and mental health in community residents. *SAGE Open, 11*(4). https://doi.org/10.1177/21582440211054477

Fava, G. A., & Bech, P. (2016). The concept of Euthymia. *Psychotherapy and Psychosomatics, 85*(1), 1–5. https://doi.org/10.1159/000441244

Fydrich, T., Sommer, G., Tydecks, S., & Brähler, E. (2009). Fragebogen zur sozialen Unterstützung (F-SozU): Normierung der Kurzform (K-14) [Social Support Questionnaire (F-SozU): Standardization of short form (K-14)]. *Zeitschrift für Medizinische Psychologie, 18*(1), 43–48.

Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German version of the satisfactionwith life scale (SWLS) psychometric properties, validity, and population-based norms. *European Journal of Psychological Assessment, 27*(2), 127–132. https://doi.org/10.1027/1015-5759/a000058

Hall, R. C. W. (1995). Global assessment of functioning: A modified scale. *Psychosomatics, 36*(3), 267–275. https://doi.org/10.1016/S0033-3182(95)71666-8

Helliwell, J. F., Layard, R., & Sachs, J. (Eds.). (2012). *World happiness report 2012*. UN Sustainable Development Solutions Network.

Hilsenroth, M. J., Ackerman, S. J., Blagys, M. D., Baumann, B. D., Baity, M. R., Smith, S. R., Price, J. L., Smith, C. L., Heindselman, T. L., Mount, M. K., & Holdwick, D. J. (2000). Reliability and validity of DSM-IV Axis V. *American Journal of Psychiatry, 157*(11), 1858–1863. https://doi.org/10.1176/appi.ajp.157.11.1858

Hirschfeld, G., von Brachel, R., & Thiele, C. (2020). Screening for health-related quality of life in children and adolescents: Optimal cut points for the KIDSCREEN-10 for epidemiological studies. *Quality of Life Research, 29*(2), 529–536. https://doi.org/10.1007/s11136-019-02324-4

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Third Edition). https://doi.org/10.1002/9781118548387

Keyes, C. L. M. (2005). Mental illness and/or mental health? Investigating axioms of the complete state model of health. *Journal of Consulting and Clinical Psychology, 73*(3), 539–548. https://doi.org/10.1037/0022-006X.73.3.539

Keyes, C. L. M., Shmotkin, D., & Ryff, C. D. (2002). Optimizing well-being: The empirical encounter of two traditions. *Journal of Personality and Social Psychology, 82*(6), 1007–1022. https://doi.org/10.1037/0022-3514.82.6.1007

Lukat, J., Margraf, J., Lutz, R., Der Veld, W. M., & Becker, E. S. (2016). Psychometric properties of the positive mental health scale (PMH-scale). *BMC Psychology, 4*(1), 1–14. https://doi.org/10.1186/s40359-016-0111-x

Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research, 46*(2), 137–155. https://doi.org/10.1023/A:1006824100041

Margraf, J., Brailovskaia, J., & Schneider, S. (2020). Behavioral measures to fight COVID-19: An 8-country study of perceived usefulness, adherence and their predictors. *PLoS One, 15*(12 December), 1–22. https://doi.org/10.1371/journal.pone.0243523

Margraf, J., Schneider, S., Soeder, U., Neumer, S., & Becker, E. S. (1996). *F-DIPS: Diagnostisches Interview bei Psychischen Störungen (Forschungsversion), Interviewleitfaden, Version 1.1, 7/96 [F-DIPS: Diagnostic interview for mental disorders (research version), interview guide, version 1.1, 7/96]*. University of Technology.

Margraf, J., Zhang, X. C., Lavallee, K. L., & Schneider, S. (2020). Longitudinal prediction of positive and negative mental health in Germany, Russia, and China. *PLoS One, 15*(6), Article e0234997. https://doi.org/10.1371/journal.pone.0234997

Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological Assessment, 5*(2), 164–172. https://doi.org/10.1037/1040-3590.5.2.164

Pavot, W., & Diener, E. (2008). The satisfaction with life scale and the emerging construct of life satisfaction. *Journal of Positive Psychology, 3*(2), 137–152. https://doi.org/10.1080/17439760701756946

R Core Team. (2022). *R: A language and environment for statistical computing*. (4.2.0). R Foundation for Statistical Computing.

Schönfeld, P., Brailovskaia, J., & Margraf, J. (2017). Positive and negative mental health across the lifespan: A cross-cultural comparison. *International Journal of Clinical and Health Psychology, 17*(3), 197–206. https://doi.org/10.1016/j.ijchp. 2017.06.003

Schumacher, J., Leppert, K., Gunzelmann, T., Strauß, B., & Brähler, E. (2005). Die Resilienzskala–Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigkeit als Personmerkmal [The resilience scale-A questionnaire to assess psychological resilience as a personal characteristic]. *Zeitschrift für Klinische Psychologie Psychiatrie und Psychotherapie, 53*(1), 16–39.

Schumacher, J., Wilz, G., Gunzelmann, T., & Brähler, E. (2000). Die Sense of Coherence Scale von Antonovsky: Teststatische Überprüfung in einer repräsentativen Bevölkerungsstichprobe und Konstruktion einer Kurzskala [Antonovsky's Sense of Coherence Scale: Statistical testing in a representative population sample and construction of a short scale]. *PPmP Psychotherapie Psychosomatik Medizinische Psychologie, 50*(12), 472–482. https://doi.org/10.1055/s-2000-9207

Söderberg, P., Tungström, S., & Armelius, B. A. (2005). Special section on the GAF: Reliability of global assessment of functioning ratings made by clinical psychiatric staff. *Psychiatric Services, 56*(4), 434–438. https://doi.org/10.1176/appi.ps.56.4.434

Suldo, S. M., & Shaffer, E. J. (2008). Looking beyond psychopathology: The dual-factor model of mental health in youth. *School Psychology Review, 37*(1), 52–68. https://doi.org/10.1080/02796015.2008.12087908

Swami, V., Stieger, S., Voracek, M., Dressler, S. G., Eisma, L., & Furnham, A. (2009). Psychometric evaluation of the tagalog and German subjective happiness scales and a cross-cultural comparison. *Social Indicators Research, 93*(2), 393–406. https://doi.org/10.1007/s11205-008-9331-7

Teismann, T., & Brailovskaia, J. (2020). Entrapment, positive psychological functioning and suicide ideation: A moderation analysis. *Clinical Psychology and Psychotherapy, 27*(1), 34–41. https://doi.org/10.1002/cpp.2403

Teismann, T., Brailovskaia, J., Siegmann, P., Nyhuis, P., Wolter, M., & Willutzki, U. (2018). Dual factor model of mental health: Co-occurrence of positive mental health and suicide ideation in inpatients and outpatients. *Psychiatry Research, 260*(May 2017), 343–345. https://doi.org/10.1016/j.psychres.2017.11.085

Tennant, R., Goens, C., Barlow, J., Day, C., & Stewart-Brown, S. (2007). A systematic review of reviews of interventions to promote mental health and prevent mental health problems in children and young people. *Journal of Public Mental Health, 6*(1), 25–32. https://doi.org/10.1108/17465729200700005

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warwick-Dinburgh mental well-being scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes, 5*, 1–13. https://doi.org/10.1186/1477-7525-5-63

Thiele, C., & Hirschfeld, G. (2021). Cutpointr: Improved estimation and validation of optimal cutpoints in r. *Journal of Statistical Software, 98*(11). https://doi.org/10.18637/jss.v098.i11

Trumpf, J., Vriends, N., Meyer, A. H., Becker, E. S., Neumer, S.-P., & Margraf, J. (2010). The Dresden Predictor Study of anxiety and depression: Objectives, design, and methods. *Social Psychiatry and Psychiatric Epidemiology, 45*(9), 853–864. https://doi.org/10.1007/s00127-009-0133-2

Tucker, K. L., Ozer, D. J., Lyubomirsky, S., & Boehm, J. K. (2006). Testing for measurement invariance in the satisfaction with life scale: A comparison of Russians and North Americans. *Social Indicators Research, 78*(2), 341–360. https://doi.org/10.1007/s11205-005-1037-5

Velten, J., Brailovskaia, J., & Margraf, J. (2021). Positive mental health scale: Validation and measurement invariance across eight Countries, genders, and age groups. *Psychological Assessment, 34*(4), 332–340. https://doi.org/10.1037/pas0001094

von Glischinski, M., von Brachel, R., & Hirschfeld, G. (2019). How depressed is "depressed"? A systematic review and diagnostic meta-analysis of optimal cut points for the Beck Depression Inventory revised (BDI-II). *Quality of Life Research, 28*(5), 1111–1118. https://doi.org/10.1007/s11136-018-2050-x

Zhang, X. C., Kuchinke, L., Woud, M. L., Velten, J., & Margraf, J. (2017). Survey method matters: Online/offline questionnaires and face-to-face or telephone interviews differ. *Computers in Human Behavior, 71*, 172–180. https://doi.org/10.1016/j.chb.2017.02.006

History
Received November 4, 2023
Revision received May 16, 2024
Accepted June 5, 2024
Published online August 8, 2024
EJPA Section / Category Positive Psychology & Assessment

Publication Ethics
The study received ethical approval from the local ethics committee of the faculty of psychology at the Ruhr University Bochum (approval number: 20110512).

Open Science
We report how we determined our sample size, all data exclusions, all data inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all measures in the study, and all analyses including all tested models. If we use inferential tests, we report exact p values, effect sizes, and 95% confidence intervals.
Open Data: I confirm that there is sufficient information for an independent researcher to reproduce all the reported results, including codebook if relevant (Bonnin et al., 2024b).
Open Materials: I confirm that there is sufficient information for an independent researcher to reproduce all the reported methodologies.
Open Analytic Code: I confirm that all the scripts, code, and outputs needed to reproduce the results are provided (Bonnin et al., 2024a).
Preregistration of Studies and Analysis Plans: This study was not preregistered.

Funding
This study was supported by the Alexander von Humboldt Professorship awarded to Jürgen Margraf. Open access publication enabled by Ruhr University Bochum.

ORCID
Gabriel Bonnin
https://orcid.org/0000-0002-7084-8774
Gerrit Hirschfeld
https://orcid.org/0000-0003-2143-4564
Ruth von Brachel
https://orcid.org/0000-0003-4117-1511

Gabriel Bonnin
Mental Health Research and Treatment Center
Faculty of Psychology
Ruhr-Universität Bochum
Massenbergstraße 9-13
44787 Bochum
Germany
gabriel.bonnin@ruhr-uni-bochum.de