



## A clinimetric analysis of the euthymia, resilience, and positive mental health scales

Danilo Carrozzino<sup>a</sup>, Kaj S. Christensen<sup>b</sup>, Giovanni Mansueto<sup>c,d</sup>, Julia Brailovskaia<sup>e</sup>, Jürgen Margraf<sup>e</sup>, Fiammetta Cosci<sup>c,f,\*</sup>

<sup>a</sup> Department of Psychology, Alma Mater University of Bologna, Bologna, Italy

<sup>b</sup> Research Unit for General Practice and Section for General Medical Practice, Department of Public Health, Aarhus University, Aarhus, Denmark

<sup>c</sup> Department of Health Sciences, University of Florence, Florence, Italy

<sup>d</sup> Department of Psychology, Sigmund Freud University, Milan, Italy

<sup>e</sup> Mental Health Research and Treatment Center, Department of Clinical Psychology and Psychotherapy, Ruhr-Universität Bochum, Bochum, Germany

<sup>f</sup> Department of Psychiatry & Neuropsychology, Maastricht University, Maastricht, The Netherlands

### ARTICLE INFO

#### Keywords:

Euthymia  
Clinimetrics  
Positive mental health  
Resilience  
Validity

### ABSTRACT

**Background:** Euthymia has been described as a transdiagnostic construct characterized by the absence of mood disturbances and the presence of an integration of psychic forces, such as coping strategy and well-being. A multidimensional measure, the Euthymia Scale (ES), has been proposed to assess it. We investigated construct and concurrent validity of the ES.

**Methods:** Construct validity was studied via Rasch and Mokken analyses and compared with that of the 11-item Resilience Scale (RS-11) and 9-item Positive Mental Health Scale (PMH-Scale). A total of 951 participants were recruited (77.5% women;  $24.86 \pm 5.62$  years).

**Results:** The ES, RS and PMH demonstrated similar sensitivity and construct validity. Findings indicate minor needs for adjustments only. As expected the ES demonstrated a strong negative correlation with neuroticism.

**Limitations:** The convenience sample of subjects recruited primarily from female Italian university students and a community-based data collection limit the generalizability of the present findings. The cross-sectional design precludes the assessment of test-retest reliability, predictive and incremental validity. Only self-report measures and a Likert version of the ES were used.

**Conclusions:** ES is the most comprehensive measure of euthymia. The RS-11 is a valid measure of a specific component of euthymia, namely subjective ability to cope with stress and empower well-being in face of life adversities. The PMH-Scale is a valid measure of overall positive mental health.

### 1. Introduction

The term euthymia has a Greek origin and derives from the combination of “*eu*”, meaning well, and “*thymós*”, which encompasses several meanings including life energy; feelings and passions; will, desire and inclination; thought and intelligence. The definition of euthymia is ascribed to Democritus according to whom a euthymic person is satisfied with what is present and available, takes little heed of people who are envied, admires and observes the lives of those who suffer and yet endure (Kahn, 1985).

The concept of euthymia was approached in the clinical realm in 1958, when Marie Jahoda denied that mental health could be usefully

defined by identifying it with the absence of a disease (Jahoda, 1958). She suggested that the absence of disease may constitute a necessary, but not sufficient, criterion for mental health (Jahoda, 1958). Nevertheless, based on current psychiatric nosography (Diagnostic and Statistical Manual of Mental Disorders, DSM), euthymia is still generally used to refer to the absence of affective disorders (e.g., Blumberg, 2012), while a positive connotation is not mentioned. Following Jahoda’s intuition, some instruments have been developed to assess euthymia (Lukat et al., 2016; Schumacher et al., 2005; Topp et al., 2015). Among them, the WHO-5 Well-Being Index (WHO-5; Topp et al., 2015), measuring subjective well-being, the 11-item Resilience Scale (RS-11; Schumacher et al., 2005), measuring psychosocial stress resilience, and the 9-item

\* Corresponding author at: Fiammetta Cosci, Department of Health Sciences, University of Florence, via di San Salvi, 12, 50135 Florence, Italy.

E-mail address: [fiammetta.cosci@unifi.it](mailto:fiammetta.cosci@unifi.it) (F. Cosci).

<https://doi.org/10.1016/j.jad.2021.07.001>

Received 22 April 2021; Received in revised form 29 June 2021; Accepted 2 July 2021

Available online 9 July 2021

0165-0327/© 2021 Elsevier B.V. All rights reserved.

version of the Positive Mental Health Scale (PMH-Scale; Lukat et al., 2016), measuring emotional, psychological, and social aspects of well-being. However, such instruments tend to cover only specific components of euthymia (Fava and Bech, 2016), implying a need for a more comprehensive measure.

In the recent years, there has been a reappraisal of euthymia as conceptualized by Jahoda (1958). Fava and his group (Fava and Bech, 2016; Fava and Guidi, 2020; Guidi and Fava, 2020) provided a comprehensive definition of euthymia conceiving it as a transdiagnostic construct characterized by the absence of mood disturbances and the presence of an integration of psychic forces, which includes psychological flexibility and well-being, resistance to stress (i.e., resilience and anxiety- or frustration-tolerance), and a subjective sense of consistency or a unifying outlook on life which guides actions and feelings for shaping the future accordingly. Fava and colleagues also provided tools for the clinical assessment of euthymia developing a self-administered scale (Fava and Bech, 2016), the Euthymia Scale (ES), and a clinician-administered measure, the Clinical Interview for Euthymia (Fava and Guidi, 2020). The ES is a multidimensional measure assessing psychological flexibility, consistency, and resilience (Fava and Bech, 2016; Carrozzino et al., 2019). It was developed according to clinimetric principles (Fava and Bech, 2016; Fava and Guidi, 2020; Guidi and Fava, 2020).

Such principles find their roots in “clinimetrics” (Feinstein, 1982; Feinstein 1983; Feinstein 1987), the science of clinical measurements (Fava et al., 2012), which has proven useful for testing measurement properties not captured by the traditional psychometric model (Fava et al., 2018). Clinimetric Patient-Reported Outcome Measures (CLIPROM) criteria have been recently introduced to challenge the traditional views of how new indices should be developed and existing measures validated (Carrozzino et al., 2021a; Cosci, 2021). According to CLIPROM criteria (Carrozzino et al., 2021a), Item Response Theory (IRT) models (i.e., Rasch and Mokken analyses) allow to evaluate the dimensionality or construct validity of self-rating scales. The clinimetric analysis of construct validity refers to the assessment of whether items of a rating scale provide unique clinical information, implies the evaluation of the extent to which items (e.g., symptoms) included in a rating scale belong to an underlying construct (i.e., clinical dimension), and includes the examination of whether the rating scale is a valid measure of the degree (e.g., severity) of the dimension (e.g., clinical condition) under assessment (Bech, 2004; Bech, 2012; Bech, 2016a; Carrozzino et al., 2020; Carrozzino et al., 2021a).

In line with CLIPROM criteria (Carrozzino et al., 2021a), the present study aimed to investigate the construct validity of the ES, which has previously been shown to be a multidimensional measure (Carrozzino et al., 2019). The ES was administered to a larger sample to test its construct validity via Rasch and Mokken analyses and to compare it with that of the RS-11 (Schumacher et al., 2005) and the PMH-Scale (Lukat et al., 2016), two widely used unidimensional patient-reported outcome measures for the evaluation of positive dimensions of mental health (see e.g., Brailovskaia et al., 2018; Lin et al., 2020; Truskauskaitė-Kunevičienė et al., 2020). Construct validity was tested to determine whether ES, RS-11, and PMH-Scale were statistically sufficient and clinically valid measures of the degree of the underlying construct that they intended to measure. The major aim was to identify the tool which best display the clinimetric property of construct validity in the process of assessment of euthymia. The concurrent validity of the ES, a clinimetric property that refers to the assessment of the extent to which a rating scale correlates with another related and previously validated assessment instrument (Carrozzino et al., 2021a), was also evaluated to investigate whether the ES is positively associated with psychometric measures, such as the RS-11 and PMH-Scale, and to examine the extent to which the ES is negatively related to the 11-item version of the Big Five Inventory (BFI-11) Neuroticism subscale (Rammstedt and John, 2007). The rationale for using measures of resilience and positive mental health is based on the transdiagnostic concept of euthymia that appears

to be characterized by such aspects (Fava and Bech, 2016; Fava and Guidi, 2020; Guidi and Fava, 2020; Zhang et al., 2021). The rationale for using a measure of personality is based on the definition of euthymia as provided by Jahoda (1958), “as a relatively constant and enduring function of personality”, and by Bech (2016b; p. 46), as a fighting spirit personality combined with aspects of psychological well-being.

## 2. Methods

### 2.1. Procedure and participants

Data were collected between the end of April and the beginning of September 2020 via an online survey. The link to the survey was sent to all institutional email addresses of the students enrolled at the University of Florence in the academic year 2019–2020 ( $n = 51,715$ ) together with an invitation to disseminate the link via email or social media among friends and acquaintances. There were no specific requirements for participation, which was voluntary and not compensated. During data collection, all participants were geographically located in Italy. This information was provided by a self-declaration at the beginning of the survey. Overall the survey was completed by 98.3% of the 952 individuals who activated the survey link.

The study was approved by the Ethical Committee of the University of Florence, Florence, Italy. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2013. All participants were provided instructions and gave informed consent to participate via an online form.

### 2.2. Measures

The Euthymia Scale is a 10-item multidimensional self-administered instrument (ES, Fava and Bech, 2016; Carrozzino et al., 2019) assessing psychological flexibility, consistency, and resilience (see appendix). It incorporates Jahoda’s (1958) conceptualization of individual balance among psychic forces. Five items were derived from the WHO-5 index (Topp et al., 2015), five items were created on purpose. Each item is scored dichotomously as 1 (*true*) or 0 (*false*). The tool was found to display the clinimetric property of incremental validity in the process of assessment of dimensions of psychological well-being (e.g., environmental mastery, purpose in life, self-acceptance) and neuroticism (Carrozzino et al., 2019). For the present study and following the aim of homogenizing the response format of the ES to the other two scales examined, the ES was presented in a Likert scale response format. A 6-point Likert scale (from 1 “strongly disagree” to 6 “strongly agree”) was chosen as quantifier, as it is proposed in other tools extensively used in clinical settings to assess psychological well-being states and dimensions (e.g., Ryff, 2014).

The 11-item Resilience Scale (RS-11; Schumacher et al., 2005) was developed as a measure of psychosocial stress resilience. Subjects respond to items such as “I usually manage one way or another”, indicating their level of agreement on a 7-point Likert scale ranging from 1 (“I disagree”) to 7 (“I agree”). Good reliability and validity indices have been shown (Schumacher et al., 2005).

The 9-item Positive Mental Health Scale (PMH-Scale; Lukat et al., 2016) is a person-centered instrument consisting of nine non-specific judgments developed to measure emotional, psychological, and social aspects of well-being, without explicitly referring to well-being theories. Targeted at general psychological functioning, subjects are asked to respond to statements such as “I am in good physical and emotional condition” indicating their agreement on a 4-point Likert scale ranging from 0 (“I do not agree”) to 3 (“I agree”). In various groups (student and different patient samples), the PMH-Scale showed a unidimensional structure, good test–retest reliability, internal consistency, convergent and discriminant validity (Lukat et al., 2016).

The 11-item version of the Big Five Inventory (BFI-11; Rammstedt and John, 2007) is a self-administered scale assessing personality dimensions according to the Five Factor Model (i.e., openness to experience, conscientiousness, extraversion, agreeableness and neuroticism). It includes two items for each personality factor, except agreeableness, which has three. Each item is scored on a 5-point Likert scale, from 1 (“strongly disagree”) to 5 (“strongly agree”). The BFI-11 showed to retain significant levels of reliability and validity (Rammstedt and John, 2007). For the present study, the Neuroticism Subscale was used (Item 4 and Item 9).

### 2.3. Statistical analyses

IRT analyses were conducted to examine the construct validity of the ES, RS-11, and PMH-Scale. The Rasch analyses were performed using Rasch Unidimensional Measurement Models (RUMM2030) software (Andrich et al., 2010) and the following properties were tested:

- 1 Overall fit to the model was evaluated using the chi-square item-trait interaction statistics (Pallant and Tennant, 2007; Tennant and Conaghan, 2007). These statistics provide a summary measure of how the rating scale conforms to the Rasch model expectations (Nielsen et al., 2017). A non-significant chi-square probability value indicates a good level of overall fit (Pallant and Tennant, 2007; Tennant and Conaghan, 2007).
- 2 Individual item and person fit: standardized fit residual values for items and subjects were examined for any indication of misfit.
- 3 Dimensionality or construct validity: to determine whether the ES, RS-11, and PMH-Scale were valid indices of the degree of the underlying dimension under assessment, Principal Component Analysis of residuals was conducted to identify the two most different subsets of items (i.e., the most positively and negatively factor-loading items on the first component). Paired *t*-tests were then performed to compare scores on the two subsets of items. If more than 5% of *t*-tests were significant, indices were not considered as unidimensional (Christensen et al., 2019; Nielsen et al., 2017).
- 4 Person Separation Reliability Index (PSI) was examined to evaluate the clinimetric sensitivity of the three rating scales, i.e., their ability to discriminate among subjects with different levels of the underlying trait under evaluation (Carrozzino et al., 2021b; Tennant and Conaghan, 2007).

Mokken analysis, which is the non-parametric version of IRT models (Bech, 2012; Mokken, 1971), was performed to further examine the dimensionality or scalability of the ES, RS-11, and PMH-Scale. The Mokken analysis was conducted using Stata statistical software, version 7 (Stata Corporation, College Station, TX). The Stata LoevH command was used to compute Loevinger’s coefficients. According to Mokken (Mokken, 1971), Loevinger’s coefficients (Loevinger, 1947) ranging from 0.30 to 0.39 are considered just acceptable, while a coefficient  $\geq 0.40$  is a clear demonstration of the scalability of the rating scale under assessment (Bech, 2012).

Finally, concurrent validity of ES was examined through correlation analyses between ES and the following measures: PMH-Scale, RS-11, and BFI-11 Neuroticism subscale.

## 3. Results

### 3.1. Sample characteristics

The sample consisted of 951 participants, 77.5% female. The mean age was  $24.86 \pm 5.62$  years (range: 19–70 years). Among them, 87% were students, 12% employees, 1.1% unemployed; 44.1% were single, 42.7% had a romantic partner, and 13.2% were married. For details on mean, standard deviation, maximum, minimum of scores of rating-scales see Table 1.

**Table 1**

Mean, standard deviation, minimum and maximum of rating scales under study total scores.

Scale	Mean	SD	Minimum	Maximum
ES	37.23	9.97	10.00	60.00
RS-11	54.89	12.06	12.00	77.00
PMH-9	16.10	6.02	0.00	27.00
BFI-11 Neuroticism Subscale	6.48	2.08	2.00	10.00

Note. ES = Euthymia Scale; RS-11 = 11-item Resilience Scale; PMH-9 = 9-item Positive Mental Health Scale;

BFI-11 = 11-item version of the Big Five Inventory.

### 3.2. Fit to the Rasch model

The initial analysis of the ES revealed a significant item-trait interaction statistic ( $\chi^2 = 351.38$ , degrees of freedom [*df*] = 90,  $p < 0.001$ ), indicating misfit to the Rasch model (Table 2, Analysis 1). Standardized fit residuals for items ( $SD = 3.72$ ) were not within acceptable limits. Standardized fit residuals for persons ( $SD = 0.97$ ) were found to be within acceptable limits. Rescoring all items, ordered response categories were achieved but without significantly improving the overall fit to the Rasch model (Table 2, Analysis 2). Even after adjusting the sample size (Table 2, Analysis 3), fit to the model was not achieved ( $\chi^2 = 135.12$ ,  $df = 90$ ,  $p = 0.001$ ). Model fit was achieved ( $\chi^2 = 70.16$ ,  $df = 81$ ,  $p = 0.80$ ) after the exclusion of the misfitting item (i.e., Item 4).

Model fit statistics for the RS-11 are reported in Table 3. The initial analysis showed a significant item-trait interaction statistic ( $\chi^2 = 273.29$ ,  $df = 99$ ,  $p < 0.001$ ), indicating misfit to the Rasch model (Table 3, Analysis 1). Standardized fit residuals for items ( $SD = 3.39$ ) were not within acceptable limits. Standardized fit residuals for persons ( $SD = 0.86$ ) were found to be within acceptable limits. Rescoring all items, ordered response categories were achieved but without significantly improving the overall fit to the Rasch model (Table 3, Analysis 2). Even after adjusting the sample size (Table 3, Analysis 3), fit to the model was not achieved ( $\chi^2 = 141.88$ ,  $df = 99$ ,  $p = 0.003$ ). Model fit was achieved ( $\chi^2 = 105.25$ ,  $df = 90$ ,  $p = 0.130$ ) after the exclusion of the misfitting item (i.e., Item 10).

Model fit statistics for the PMH-Scale are reported in Table 4. The initial analysis of the PMH-Scale revealed a significant item-trait interaction statistic ( $\chi^2 = 340.73$ ,  $df = 81$ ,  $p < 0.001$ ), indicating misfit to the Rasch model (Table 4, Analysis 1). Standardized fit residuals for items ( $SD = 5.21$ ) were not within acceptable limits. Standardized fit residuals for persons ( $SD = 1.91$ ) were found to be within acceptable limits. Even after adjusting the sample size (Table 4, Analysis 2), fit to the Rasch model was not achieved ( $\chi^2 = 184.38$ ,  $df = 81$ ,  $p < 0.001$ ). Model fit was achieved ( $\chi^2 = 80.65$ ,  $df = 64$ ,  $p = 0.08$ ) after the exclusion of the misfitting item (i.e., Item 9).

### 3.3. Dimensionality and scalability

Testing for dimensionality revealed significant *t*-tests outside the critical value of 5%, suggesting that the ES was not unidimensional (Table 2, Analysis 1–4). Using the Mokken analysis, the total score of the ES had an acceptable scalability (Loevinger’s coefficient of homogeneity of 0.47).

As to the RS-11, paired *t*-tests comparisons showed that more than 5% of *t*-tests were significant, suggesting that this rating scale was not unidimensional (Table 3, Analysis 1–4). Using the Mokken analysis, the total score of the RS-11 had an acceptable scalability (Loevinger’s coefficient of homogeneity of 0.44). The exclusion of the misfitting item (“Sometimes I make myself do things whether I want to or not”) further improved the dimensionality or scalability (Loevinger’s coefficient of homogeneity of 0.47) of the RS-11.

Concerning the PMH-Scale, even after adjusting the sample size, more than 5% of *t*-tests were significant, suggesting that this measure

**Table 2**  
Model fit statistics for Euthymia Scale (data were available for 947 over 951 participants).

Action	Analysis	Model fit (overall)	Item fit residual, mean (SD)	Person fit residual, mean (SD)	PSI	Dimensionality, significant t-tests (%)
Original sample	1	$\chi^2(90) = 351.38, p < 0.001$	0.22 (3.72)	0.16 (0.97)	0.88	10.30
Rescoring all items (011223)	2	$\chi^2(90) = 255.91, p < 0.001$	-0.38 (3.45)	0.24 (1.46)	0.87	11.36
Adjusted sample (n=500)	3	$\chi^2(90) = 135.12, p = 0.001$	-0.38 (3.45)	0.24 (1.46)	0.87	11.36
Delete item 4	4	$\chi^2(81) = 70.16, p = 0.80$	-0.22 (3.22)	0.09 (1.59)	0.87	10.20

Note.  $\chi^2$  = chi-square; p = probability; SD = standard deviation; PSI = person separation index (with extremes).

**Table 3**  
Model fit statistics for Resilience scale (data were available for 944 over 951 participants).

Action	Analysis	Model fit (overall)	Item fit residual, mean (SD)	Person fit residual, mean (SD)	PSI	Dimensionality, significant t-tests (%)
Original Sample	1	$\chi^2(99) = 273.29, p < 0.001$	0.26 (3.39)	0.61 (0.86)	0.87	7.89
Rescoring All Items (0112345)	2	$\chi^2(99) = 267.87, p < 0.001$	0.65 (3.52)	0.59 (0.93)	0.88	7.78
Adjusted Sample (N=500)	3	$\chi^2(99) = 141.88, p = 0.003$	0.65 (3.52)	0.59 (0.93)	0.88	7.78
Delete Item 10	4	$\chi^2(90) = 105.25, p = 0.130$	0.61 (2.52)	0.61 (1.01)	0.88	8.52

Note.  $\chi^2$  = chi-square; p: probability; SD = standard deviation; PSI = person separation index (with extremes).

**Table 4**  
Model fit statistics for Positive Mental Health scale (data were available for 924 over 951 participants).

Action	Analysis	Model fit (overall)	Item fit residual, mean (SD)	Person fit residual, mean (SD)	PSI	Dimensionality, significant t-tests (%)
Original sample	1	$\chi^2(81) = 340.73, p < 0.001$	-0.65 (5.21)	0.71 (1.91)	0.90	7.05
Adjusted sample (n=500)	2	$\chi^2(81) = 184.38, p < 0.001$	-0.65 (5.21)	0.71 (1.91)	0.90	7.05
Delete item 9	4	$\chi^2(64) = 80.65, p = 0.08$	-0.59 (3.47)	0.81 (2.19)	0.90	3.89

Note.  $\chi^2$  = chi-square; p = probability; SD = standard deviation; PSI = person separation index (with extremes).

was not unidimensional (Table 4, Analysis 1-2). After the exclusion of the misfitting item (“I am a calm, balanced human being”), less than 5% of t-tests were significant, indicating that the 8-item version of the PMH-Scale was unidimensional (Table 4, Analysis 4). The dimensionality of the total score of the 8-item version of the PMH-Scale was further confirmed by the Mokken analysis, which showed a Loevinger’s coefficient of scalability of 0.65.

### 3.4. Local dependency

Local dependency was investigated evaluating whether the response to one item was dependent on the response to another item after controlling for the underlying trait. Local dependency was detected between Items 1 (“If I become sad, anxious or angry it is for a short time”) and 8 (“I generally feel active and vigorous”) of the ES. As to the RS-11, the following item-pairs were found to be locally dependent: Items 1 (“When I make plans I follow through with them”) and 8 (“I can usually find something to laugh about”), Items 1 (“When I make plans I follow through with them”) and 9 (“I can usually look at a situation in a number of ways”). Local dependency was also detected between Items 2 (“I enjoy my life”) and 3 (“All in all, I am satisfied with my life”) of the PMH-Scale.

### 3.5. Differential Item Functioning (DIF)

DIF was evaluated testing whether a certain form of item bias can occur when different groups of subjects (e.g., males and females) respond differently to an item despite equal levels of the underlying

trait. The item 5 of the ES (“Most of the time I can handle stress”) showed a uniform DIF for age with adults reporting higher scores. There were no indications of uniform or non-uniform DIF for the other scales.

### 3.6. Person separation reliability index (PSI)

PSI was 0.88 (Table 2, Analysis 1), indicating that the ES could reliably discriminate between respondents displaying different levels of the underlying trait under examination. PSI indices of the RS-11 were found to range from 0.87 to 0.88 (Table 3, Analysis 1-4), suggesting that this scale could reliably discriminate between subjects with different levels of the underlying construct under assessment. As to the PMH-Scale, PSI was 0.90 (Table 4, Analysis 1), indicating that this measure could sensitively discriminate between respondents displaying different levels of the underlying trait under evaluation.

### 3.7. Correlation analyses

The ES was found to be significantly and positively correlated with the RS-11 ( $r = 0.74$ ) and the PMH-Scale ( $r = 0.80$ ). A statistically significant and negative correlation was found between the ES and the BFI-11 Neuroticism subscale ( $r = -0.56$ ).

## 4. Discussion

This is the first study combining Rasch and Mokken analyses to evaluate the clinimetric properties of self-administered rating scales for

the assessment of euthymia, resilience, and positive mental health. The findings of the present study indicate that ES, RS-11, and PMH-Scale are highly sensitive and valid patient-reported outcome measures assessing different aspects of euthymia. They also entail different clinimetric properties.

The ES displayed the clinimetric features of sensitivity and construct validity. It sensitively discriminated between subjects with different levels of euthymia and its total score was found to be a comprehensive and valid measure of the degree of euthymia. This is in line with Zhang et al. (2021) who showed that the ES is a highly sensitive clinimetric index that can be used as a screening measure to detect vulnerability to depression. However, caution should be paid when using the Likert version of the ES since it was found to include two locally dependent items.

The same problem of local dependency was observed with the RS-11, which was found to include inter-correlated items providing clinically redundant information. The findings of the present study also indicate that it is preferable to use a 10-item version of the RS-11 that does not include the misfitting item “Sometimes I make myself do things whether I want to or not”. Such a brief version appears to cover a specific component of euthymia (Fava and Bech, 2016), since it applies to the assessment of the degree of individual resilience, a stable personality characteristic that enhances individual adaptation (Fletcher and Sarkar, 2013; Kocalevent et al., 2015; Rutter, 1993; Wagnild and Collins, 2009; Wagnild and Young, 1993).

The PMH-Scale, particularly the 8-item version that does not include the misfitting item “I am a calm, balanced human being”, was found to be a valid measure of the degree of positive mental health. This is in line with Lukat et al. (2016) who suggested to use this scale to provide “a quick overall assessment of positive mental health”. It should be noted, however, that the PMH-Scale does not primarily focus on Jahoda’s component of psychological flexibility that is an important core feature of positive mental health (Kashdan and Rottenberg, 2010). Notably, such items are included in the ES. Therefore, the use of ES might be of benefit in studies that specifically focus on psychological flexibility.

The present findings also indicate that the ES entails the clinimetric property of concurrent validity. Positive associations were found between ES and measures of resistance to stress (RS-11) and positive mental health (PMH-Scale). This result is consistent with the concept of euthymia, a multidimensional construct characterized by aspects of resilience combined with the subjective ability of feeling cheerful, calm, relaxed, active, vigorous, and interested in things (Fava and Bech, 2016). It should be noted, however, that in clinimetrics a high correlation does not indicate similar clinical validity: rating scales may have a common content that ensures a positive association, but they may display differential validity (Carrozzino et al., 2021a). In addition, correlation coefficients are often of statistical but not of clinical significance (Carrozzino et al., 2021a). Caution should be, therefore, paid when interpreting these results. Further, a negative relationship was detected between ES and neuroticism, providing additional evidence that the ES is a comprehensive measure of euthymia covering both personality traits of emotional stability/balance and factors of psychological well-being (Bech et al., 2016; Carrozzino et al., 2016). This is consistent with Jahoda (1958), who defined euthymia “as a relatively constant and enduring function of personality”, and with Bech (2016b; p. 46), who noted that euthymia can be defined as a fighting spirit personality combined with aspects of psychological well-being. This finding can be also interpreted based on the negative relationship between euthymia and dysthymia (Bech et al., 2016; Carrozzino et al., 2016; Zhang et al., 2021). Dysthymia and euthymia should not be considered as opposite ends of the same scale since, from a clinimetric point of view, they measure two different dimensions, namely negative mental health or dysthymia and positive mental health or euthymia (Carrozzino et al., 2016).

## 5. Limitations

The present study has some limitations. The convenience sample of subjects recruited primarily from predominantly female Italian university students limits the generalizability of findings. Future studies should include subjects of different age and education enrolled from the general population. A community-based sample was used. Considering that clinimetrics and CLIPROM criteria apply to the assessment of a number of issues that include both the evaluation of clinical phenomena and the assessment of positive dimensions of mental health, this is a relative limitation which, however, encourages future research in clinical samples (e.g., depressed patients). A cross-sectional design was used, precluding the assessment of test-retest reliability, predictive and incremental validity. Studies using a longitudinal design are needed. Only self-report measures were used. Future research, making use of the clinical judgment of experienced clinicians or including clinician-rated scales, are recommended particularly to test the clinical validity of the ES, RS-11, and PMH-Scale. Finally, future studies using the original version of ES (Fava and Bech, 2016) are needed to further assess its clinimetric properties.

## 6. Conclusions

The concept of euthymia is worthy of being reappraised in the clinical realm when mental health is the issue. Currently, several clinimetric indices are available to assess it. The present findings reveal that the 10-item version of the RS-11 might evaluate a specific component of euthymia, namely the subjective ability to cope with stress and empower well-being in face of life adversities (i.e., resilience). The 8-item version of the PMH-Scale might provide an overall evaluation of the degree of positive mental health. When a more comprehensive evaluation is needed, the ES should be preferred in the process of assessment of the degree of euthymia. These tools are in need of being implemented in clinics and research for comprehensively assessing patients and verifying whether treatments (both pharmacological and non-pharmacological) are effective also in building mental health. As Marie Jahoda suggested, “let us define mental health as the adjustment of human beings to the world and to each other with a maximum of effectiveness and happiness. Not just efficiency [thus, functioning], or just contentment [thus, quality of life] – or the grace of obeying the rules of the game cheerfully [for instance, compliance]. It is all of these together. It is the ability to maintain an even temper, an alert intelligence, socially considerate behavior, and a happy dimension. This [...] is a healthy mind” (Jahoda, 1958; p. 18).

## Statements

### Author agreement

All authors have seen and approved the final version of the manuscript being submitted. They warrant that the article is the authors’ original work, hasn’t received prior publication and isn’t under consideration for publication elsewhere.

### Funding source declaration

No funding or research grants (and their source) were received.

### Contributors

Danilo Carrozzino wrote the first draft of the manuscript; Kaj S. Christensen conducted the statistical analyses; Giovanni Mansueto revised the final version of the paper; Julia Brailovskaia revised the final version of the paper, Jürgen Margraf revised the final version of the paper; Fiammetta Cosci projected the research, supervised data collection, collected data, revised all versions of the paper.

All authors approved the final article.

### Funding

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

### Data accessibility

Data analysed in this study are available from the corresponding author.

### Declaration of Competing Interest

None.

### Acknowledgements

No acknowledgements.

### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jad.2021.07.001.

### References

- Andrich, D., Lyne, A., Sheridan, B., Luo, G., 2010. RUMM 2030. RUMM Laboratory, Perth, Australia.
- Bech, P., 2004. Modern psychometrics in clinimetrics. *Psychother. Psychosom.* 73 (3), 134–138. <https://doi.org/10.1159/000076448>.
- Bech, P., 2012. *Clinical Psychometrics*. Wiley, Blackwell, Oxford.
- Bech, P., 2016a. Clinimetric dilemmas in outcome scales for mental disorders. *Psychother. Psychosom.* 85 (6), 323–326. <https://doi.org/10.1159/000448810>.
- Bech, P., 2016b. Measurement-based care in mental disorders. Springer-Verlag, New York.
- Bech, P., Carrozzino, D., Austin, S.F., Møller, S.B., Vassend, O., 2016. Measuring euthymia within the neuroticism scale from the NEO personality inventory: a mokken analysis of the norwegian general population study for scalability. *J. Affect. Disord.* 193, 99–102. <https://doi.org/10.1016/j.jad.2015.12.039>.
- Blumberg, H.P., 2012. Euthymia, depression, and mania: what do we know about the switch? *Biol. Psychiatry*. 71 (7), 570–571. <https://doi.org/10.1016/j.biopsych.2012.02.003>.
- Brailovskaia, J., Teismann, T., Margraf, J., 2018. Cyberbullying, Positive Mental Health and Suicide Ideation/Behavior. *Psychiatry. Res.* 267, 240–242. <https://doi.org/10.1016/j.psychres.2018.05.074>.
- Carrozzino, D., Christensen, K.S., Cosci, F., 2021b. Construct and criterion validity of patient-reported outcomes (PROs) for depression: a clinimetric comparison. *J. Affect. Disord.* 283, 30–35. <https://doi.org/10.1016/j.jad.2021.01.043>.
- Carrozzino, D., Patierno, C., Fava, G.A., Guidi, J., 2020. The Hamilton Rating Scales for Depression: a critical review of clinimetric properties of different versions. *Psychother. Psychosom.* 89 (3), 133–150. <https://doi.org/10.1159/000506879>.
- Carrozzino, D., Patierno, C., Guidi, J., Berrocal Montiel, C., Cao, J., Charlson, M.E., Christensen, K.S., Concato, J., De las Cuevas, C., de Leon, J., Eöry, A., Fleck, M.P., Furukawa, T.A., Horwitz, R.I., Nierenberg, A.A., Rafanelli, C., Wang, H., Wise, T.N., Sonino, N., Fava, G.A., 2021a. Clinimetric Criteria for Patient-Reported Outcome Measures. *Psychother. Psychosom.* 90, 222–232. <https://doi.org/10.1159/000516599>.
- Carrozzino, D., Svicher, A., Patierno, C., Berrocal, C., Cosci, F., 2019. The Euthymia Scale: A Clinimetric Analysis. *Psychother. Psychosom.* 88 (2), 119–121. <https://doi.org/10.1159/000496230>.
- Carrozzino, D., Vassend, O., Bjørndal, F., Pignolo, C., Olsen, L.R., Bech, P., 2016. A clinimetric analysis of the Hopkins Symptom Checklist (SCL-90-R) in general population studies (Denmark, Norway, and Italy). *Nord. J. Psychiatry.* 70 (5), 374–379. <https://doi.org/10.3109/08039488.2016.1155235>.
- Christensen, K.S., Oerboel, E., Nielsen, M.G., Bech, P., 2019. Diagnosing depression in primary care: a rasch analysis of the major depression inventory. *Scand. J. Prim. Health. Care.* 37 (1), 256–263. <https://doi.org/10.1080/02813432.2019.1568703>.
- Cosci, F., 2021. Clinimetric Perspectives in Clinical Psychology and Psychiatry. *Psychother. Psychosom.* 90 (4), 217–221. doi: 10.1159/000517028.
- Fava, G.A., Bech, P., 2016. The concept of euthymia. *Psychother. Psychosom.* 85 (1), 1–5. <https://doi.org/10.1159/000441244>.
- Fava, G.A., Carrozzino, D., Lindberg, L., Tomba, E., 2018. The clinimetric approach to psychological assessment: a tribute to Per Bech, MD (1942–2018). *Psychother. Psychosom.* 87 (6), 321–326. <https://doi.org/10.1159/000493746>.
- Fava, G.A., Guidi, J., 2020. The pursuit of euthymia. *World. Psychiatry.* 19 (1), 40–50. <https://doi.org/10.1002/wps.20698>.
- Fava, G.A., Tomba, E., Sonino, N., 2012. Clinimetrics: the science of clinical measurements. *Int. J. Clin. Pract.* 66 (1), 11–15. <https://doi.org/10.1111/j.1742-1241.2011.02825.x>.
- Feinstein, A.R., 1982. T. Duckett Jones memorial lecture. The Jones criteria and the challenge of clinimetrics. *Circulation.* 66 (1), 1–5. <https://doi.org/10.1161/01.CIR.66.1.1>.
- Feinstein, A.R., 1983. An additional basic science for clinical medicine: IV. The development of clinimetrics. *Ann. Intern. Med.* 99 (6), 843–848. <https://doi.org/10.7326/0003-4819-99-6-843>.
- Feinstein, A.R., 1987. *Clinimetrics*. Yale University Press, New Haven.
- Fletcher, D., Sarkar, M., 2013. Psychological resilience. A Review and critique of definitions, concepts, and theory. *Eur. Psychol.* 18 (1), 12–23. <https://doi.org/10.1027/1016-9040/a000124>.
- Guidi, J., Fava, G.A., 2020. The emerging role of euthymia in psychotherapy research and practice. *Clin. Psychol. Rev.* 82, 101941. <https://doi.org/10.1016/j.cpr.2020.101941>.
- Jahoda, M., 1958. *Current Concepts of Positive Mental Health*. Basic Books, New York. <https://doi.org/10.1037/11258-000>.
- Kashdan, T.B., Rottenberg, J., 2010. Psychological flexibility as a fundamental aspect of health. *Clin. Psychol. Rev.* 30 (7), 865–878. <https://doi.org/10.1016/j.cpr.2010.03.001>.
- Kahn, C.H., 1985. Democritus and the origins of moral psychology. *Am. J. Philol.* 106 (1), 1–31. <https://doi.org/10.2307/295049>.
- Kocalevent, R.D., Zenger, M., Heinen, I., Dwinger, S., Decker, O., Brähler, E., 2015. Resilience in the general population: standardization of the resilience scale (RS-11). *PLoS. One.* 10 (11) <https://doi.org/10.1371/journal.pone.0140322>. Article e0140322.
- Lin, M., Wolke, D., Schneider, S., Margraf, J., 2020. Bullying history and mental health in university students: the mediator roles of social support, personal resilience, and self-efficacy. *Front. Psychiatry.* 10, 960. <https://doi.org/10.3389/fpsy.2019.00960>.
- Loevinger, J., 1947. A systematic approach to the construction and evaluation of tests of ability. *Psychol. Monogr.* 61 (4) <https://doi.org/10.1037/h0093565> i-49.
- Lukat, J., Margraf, J., Lutz, R., van der Veld, W.M., Becker, E.S., 2016. Psychometric properties of the positive mental health scale (PMH-scale). *BMC. Psychology.* 4, 8. <https://doi.org/10.1186/s40359-016-0111-x>.
- Mokken, R.J., 1971. A theory and procedure of scale analysis. De Gruyter Mouton, The Netherlands. <https://doi.org/10.1515/9783110813203>.
- Nielsen, M.G., Ørnboel, E., Vestergaard, M., Bech, P., Christensen, K.S., 2017. The construct validity of the major depression inventory: a Rasch analysis of a self-rating scale in primary care. *J. Psychosom. Res.* 97, 70–81. <https://doi.org/10.1016/j.jpsychores.2017.04.001>.
- Pallant, J.F., Tennant, A., 2007. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *Br. J. Clin. Psychol.* 46 (1), 1–18. <https://doi.org/10.1348/014466506X96931>.
- Rammstedt, B., John, O.P., 2007. Measuring personality in one minute or less: a 10-item short version of the Big Five Inventory in English and German. *J. Res. Pers.* 41 (1), 203–212. <https://doi.org/10.1016/j.jrp.2006.02.001>.
- Rutter, M., 1993. Resilience: some conceptual considerations. *J. Adolesc. Health.* 14 (8), 626–631. [https://doi.org/10.1016/1054-139X\(93\)90196-V](https://doi.org/10.1016/1054-139X(93)90196-V).
- Ryff, C.D., 2014. Psychological well-being revisited. *Psychother. Psychosom.* 83 (1), 10–28. <https://doi.org/10.1159/000353263>.
- 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 resilience as a personality characteristic]. *Z. Klin. Psychol. Psychopathol. Psychother.* 53 (1), 16–39.
- Tennant, A., Conaghan, P.G., 2007. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis. Care. Res.* 57 (8), 1358–1362. <https://doi.org/10.1002/art.23108>.
- Topp, C.W., Østergaard, S.D., Søndergaard, S., Bech, P., 2015. The WHO-5 well-being index: a systematic review of the literature. *Psychother. Psychosom.* 84 (3), 167–176. <https://doi.org/10.1159/000376585>.
- Truskauskaitė-Kuneviciene, I., Kazlauskas, E., Ostreikaite-Jureviciene, R., Brailovskaia, J., Margraf, J., 2020. Positive mental health and adjustment following life-stressors among young adults. *Curr. Psychol.* 1–6. <https://doi.org/10.1007/s12144-020-00714-3>.
- Wagnild, G.M., Collins, J.A., 2009. Assessing resilience. *J. Psychosoc. Nurs. Ment. Health. Serv.* 47 (12), 28–33. <https://doi.org/10.3928/02793695-20091103-01>.
- Wagnild, G.M., Young, H.M., 1993. Development and psychometric evaluation of the Resilience Scale. *J. Nur. Meas.* 1 (2), 165–178.
- Zhang, Y., Wang, X., Carrozzino, D., Yuan, Y., 2021. Clinimetric properties of the Chinese version of the Euthymia Scale. *Clin. Psychol. Psychother.* <https://doi.org/10.1002/cpp.2592>.