Creation and validation of the pictorial ecological momentary well-being instrument (EMOWI) for adolescents

Marie Buzzi[1] · Laetitia Minary[1] · Yan Kestens[2,3] · Nelly Agrinier[1,4] · Laetitia Ricci[4] · Jonathan Epstein[1,4]

Abstract

Purpose Adolescence is characterized by the ongoing maturation of emotion-regulation skills and increased emotional reactivity. There is a need for a measurement tool suitable to the Ecological Momentary Assessment methodology, to better capture within-day variations in well-being, and provide fine-grained data that can help understand how environments, behaviors, and health intersect. This paper presents the development and evaluation of the Ecological MOmentary Well-Being Instrument for adolescents, designed for use in EMA.

Methods A mixed-methods study was conducted, using both qualitative and quantitative approaches, to develop and assess the EMOWI. A literature review, pictorial production by graphic designers, and qualitative interviews with French and Canadian professionals and adolescents helped design and evaluate the scale face validity. Quantitative evaluation of dimensionality, reliability, and validity was conducted in two samples of French 8th graders.

Results The resulting 8-item EMOWI showed excellent face validity. Confirmatory factor analysis supported a single factor hypothesis (RMSEA = 0.072). Internal consistency (Cronbach's alpha = 0.85) and intraday test–retest reliability (ICC = 0.83) were high. Correlations with existing scales were consistent with preset hypotheses. Ceiling effects were evidenced for all items, yet not on the global score. Quantitative estimations were similar for the verbal and pictorial versions, but qualitative findings argued in favor of the pictorial version.

Conclusion The 8-item pictorial EMOWI is a short and innovative instrument to measure momentary well-being in adolescents aged 12 to 17 years. Its strong psychometric properties and its acceptability among adolescents make it an excellent candidate instrument for the Ecological Momentary Assessment of well-being in this population.

Keywords Well-being · Adolescent · Mental health · Ecological momentary assessment · Patient-Reported outcomes

Plain English summary

There is a lack of measurement tools to evaluate momentary well-being among adolescents, allowing to account for the intense emotional reactivity characteristic of this population. This paper presents a new instrument, the pictorial Ecological MOmentary Well-Being Instrument (EMOWI) for adolescents. The EMOWI uses sliders and pictorial anchors, designed for easiness of use on a screen and easy understanding, to be answered several times a day. The aim of this study was to report the development and psychometric properties of the EMOWI, and compare them to the properties of a verbal version of the same scale. Both the verbal and pictorial scales had excellent psychometric properties, although adolescents preferred the pictorial version. The EMOWI could help researchers, clinicians, and decision-makers in healthcare assess intraday variations in adolescents' mental well-being, and support research on how environmental conditions, dedicated mental health interventions, and health behaviors interact.

✉ Marie Buzzi
m.buzzi@chru-nancy.fr

1 Université de Lorraine, APEMAC, 54000 Nancy, France

2 Centre de Recherche en Santé Publique (CReSP), 7101, rue du Parc, Montréal, QC H3N 1X9, Canada

3 École de Santé Publique de l'Université de Montréal (ESPUM), 7101, rue du Parc, Montréal, QC H3N 1X9, Canada

4 CHRU-Nancy, INSERM, Université de Lorraine, CIC, Epidémiologie Clinique, 54000 Nancy, France

Introduction

Mental well-being is a broad and complex construct, defined by the World Health Organization as a state in which "an individual realizes his own abilities, can cope with the normal stresses of life, can work productively and is able to make a contribution to his community" [1]. While an important health outcome by itself, its association with various risk behaviors during adolescence, such as substance abuse or sedentary living, has been repeatedly documented in literature [2–4], making it a key factor to consider when promoting adolescents' health. To this day, mental well-being has been mostly considered and measured as a stable construct, with dedicated scales using long or non-specific timeframes [5–9]. However, adolescence is characterized by the ongoing maturation of emotion-regulation skills, resulting in intense emotional reactivity [10, 11]. Ecological Momentary Assessment (EMA), which consists of repeatedly collecting the immediate or recent states and experiences of subjects in their natural environment [12], offers important potential to capture the fluctuation of mental well-being in adolescents. Yet, none of the existing scales of well-being are suitable for EMA, either because they are too long to be administered repeatedly; or because the dimensions they measure are not relevant for momentary assessment. Many authors aiming to assess mental well-being with EMA have thus decided to focus on its emotional dimension [13–16]. However, this solution leaves out the psychological and social aspects of well-being, which have been identified as decisive by several authors [17–20], and may be especially important in adolescents because of the intense changes occurring during this period of life [21]. We report on the development and psychometric properties of the Ecological MOmentary Well-Being Instrument (EMOWI), a new scale of mental well-being tailored for EMA use, that includes key dimensions of adolescents' mental well-being, as identified in the literature. We were initially inspired by the Affective Slider, a self-assessment scale designed to measure immediate emotions [22] and composed of two sliders whose ends are represented by stylized facial expressions. Pictorial scales such as the Affective Slider have several benefits, as compared to verbal scales [23]: since there is no "necessity for translating feelings into words" [24], they tend to be more intuitively comprehensible, especially in the context of repeated measurement [25]. Furthermore, the target population being adolescents, we assumed a pictorial scale would be more engaging, and increase motivation in respondents [23, 26].

Methods

A mixed-methods study was conducted, following the scale development process defined by Boateng et al. [27], combined with Sauer et al.'s approach to pictorial scales development [23] (Fig. 1). We did not engage in an actual EMA validation design at this stage, considering the need to first evaluate the instruments' basic properties in terms of acceptability, dimensionality, reliability, and content validity [28].

Phase 1: item development

Identification of domain(s) and item generation

Many definitions of mental well-being involve one of two distinct dimensions: (1) emotional or subjective well-being, from a hedonic perspective, which covers people's "emotional responses, domain satisfactions, and global judgments of life satisfaction." [29]; and (2) psychological well-being, from an eudemonic perspective, which refers to cognitive functioning and self-realization [30] (e.g., self-esteem, resilience, autonomy, etc.). While scales intending to measure mental well-being usually focus on one of these two perspectives [6, 8, 31, 32], there is growing evidence that both are complementary, and that together, they can help grasp the complexity of well-being in adults [20] and adolescents [33, 34]. The choice was therefore to not rely on a predefined theoretical framework, but rather, identify and include in the scale important dimensions related to adolescents' mental
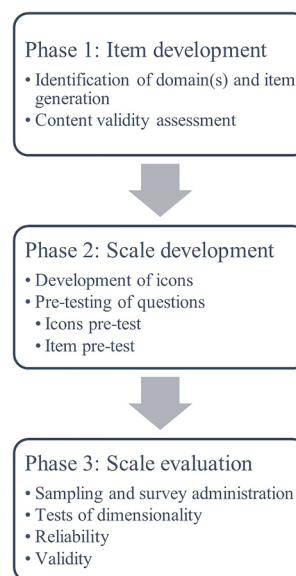

Fig. 1 Scale development process framework for the EMOWI, derived from recommendations by Boateng et al. [27], and Sauer et al. [23]

well-being, whether affective or cognitive, to capture a wide and universal conception of this construct.

We conducted a literature review using the PubMed database and the following terms: "((("well-being"[All Fields]) OR ("wellbeing"[All Fields])) AND (("adolescent"[All Fields]) OR ("youth"[All Fields]) OR ("teenager"[All Fields])))". We searched for publications in English or French language with no date restriction and made use of the backward reference searching method [35] to identify further publications. For each identified dimension, definitions and related scales items were compiled.

We then reviewed the identified dimensions to retain only those that met two conditions: perceived suitability to the adolescent population, and potential for variability in time, in line with the momentary assessment framework. This selection process resulted in a first set of items, based on existing scales and definitions. The scale was meant to be self-administered and response categories were represented by a line segment between two statements (later named "anchors" in this paper), corresponding to the ends of each dimension, and formulated according to a consensus within the development team (NA, MB, JE, YK, LM). Items were reviewed for appropriate wording, and specific attention was paid to ensure transcultural understanding in both France and French-speaking Canada.

Content validity assessment

To evaluate relevance, representativeness, and item quality, this first draft was submitted to a panel of experts, composed of a 14-year-old Quebecer adolescent, a Canadian mental health epidemiologist, and three French professionals specialized in adolescent psychiatry or psychology, using a written questionnaire. Answers were imported into NVivo Software v11, and a thematic analysis was conducted using Tourangeau's theory of cognitive processes [36]. As evidenced by Tourangeau and Willis [37], four cognitive tasks are required from respondents when completing a questionnaire (comprehension, retrieval, decision, and response), the evaluation of which may help identify sources of error in answers. As the "retrieval" process (*i.e.,* recallability and recall strategies used by respondents to retrieve relevant information from their memory) is not relevant in momentary measurement, we used instead the concept of "temporal comprehension", defined by Murphy et al. [38] as "the extent to which the respondent understands that the question is referring to the current period". Data were independently coded into Tourangeau's categories by MB and LR, and coding was compared using Cohen's kappa coefficients, with a desired value > 0.80 [39].

Phase 2: scale development

Development of icons

Professional graphic designers from Atelier Valmy (Paris, France) were asked to produce icons to represent each anchor, that could be understandable by adolescents as young as 12 years old and would be as universally comprehensible as possible. This process was iterative, with designers adjusting their work according to comments from the development team.

Pre-testing of questions

Icons pre-test Using the think aloud technique [40], pictorial anchors were submitted to a panel of eight adolescents aged 12 to 17 and recruited through convenience sampling, with careful attention paid to recruitment to ensure gender, age, and geographic diversity. Two were Quebecers and six were French. Individual interviews were conducted by MB through videoconference in January 2021 and lasted 30 min each. First, participants were asked to give their interpretation of each pair of icons without any verbal clue. They were then informed about the underlying concept and were asked to report their suggestions to improve the icons, as well as their preferences when several versions were available for a same concept. Interviews were recorded and transcribed verbatim. Data were imported into Iramuteq v0.7 and lemmatized [41] to identify the terms adolescents most frequently used to describe each anchor. These terms were then compared to the definitions retrieved from the literature, to assess the extent to which participants' understanding was consistent with the intended meaning of icons. As recommended by Sauer et al. [23], this process was iterative, with icons being refined according to adolescents' suggestions after the first four interviews.

Contrary to the Affective Slider, and because some of the targeted concepts were considered as complex to understand, a verbal instruction was kept above each item to reduce misunderstanding. Two versions of the scale were created for comparison: one with pictorial anchors (icons), and one with verbal anchors, later named "pictorial" and "verbal" versions, respectively.

Item pre-test To identify any remaining source of measurement error [38] and assess face validity, the scale was submitted to a new sample of seven French adolescents aged 12 to 16 years. Individual cognitive interviews were conducted by MB in April 2021, using the verbal probing technique [42], and lasted 40 min each. For each item, participants were encouraged to share any suggestion to improve understanding and usability before quantitative testing. They were also questioned about their opinion regarding the respec-

tive benefits and drawbacks of verbal and pictorial anchors. Interviews were recorded and transcribed verbatim. Data were imported into NVivo software, and Tourangeau's model of cognitive processes [36] was used to conduct a thematic analysis. Data were independently coded into Tourangeau's categories by MB and LR, with the addition of a new category called "pictorial comprehension", to account for participants' understanding of icons, and coding was compared using Cohen's kappa coefficients.

Phase 3: scale evaluation

Sampling and survey administration

Evaluation of the scale psychometric properties was conducted as part of a pilot study for the EXIST project, using data collected in May and June 2021 in a sample of 8th graders from four French schools located in the French Grand Est region. Schools were selected in collaboration with the regional board of education, according to their characteristics (rural or urban location, socioeconomic environment, and size), to represent a diverse population of adolescents. As recommended by the COSMIN Study Design checklist for Patient-reported outcome measurement instruments [43], the aim was to collect at least 200 observations.

Both versions of the scale (pictorial and verbal) were administered twice on the same school day during class hours, first in the morning and then in the afternoon, to investigate intraday test–retest reliability, given the momentary nature of measurement. On the first administration, sociodemographic information was collected: age (years), and gender (female, male, other), and a unique anonymous identifier was given to each student to allow matching of the two data points.

The scale was implemented in the Research Suite, an online platform provided by Polygon Research Inc., and administered on numeric tablets. The slider was presented on a continuous line with 11 steps. Participants were asked to move it to the position that best described their experience at the moment they were answering. The position was interpreted for each item as a discrete score ranging from 0 to 10, with the middle point at 5, where the departure position was located. To go to the next question, participants had to at least click once on the slider, each item being presented on a separate page. A global score was calculated by totaling the score for all items, with reverse scoring of the 2nd, 3rd, and 4th items. The use of single items was not considered.

Sample characteristics were summarized using descriptive statistics.

Tests of dimensionality

Despite the various dimensions of mental well-being, the hypothesis was that the scale would observe a unidimensional structure, with well-being as the higher factor. A confirmatory factor analysis (CFA) was carried out to test the appropriateness of this hypothesis, first on a training set, then, after modification and retest of the model if needed, on a validation set, both sets being independent subsamples of the whole sample. The comparative fit index (CFI) and Tucker–Lewis index (TLI) were assessed with a desired level of > 0.9. The Root Mean Square Error of Approximation (RMSEA) was calculated, as well as the items loadings, with a desired value of < 0.08 and > 0.3, respectively [27].

Reliability

Internal consistency Cronbach's alpha was calculated to measure internal consistency, with a desired value between 0.70 and 0.95 [44].

Test–retest reliability To determine whether a change in the global score was expected between the two measurement points, participants were asked an additional question at the beginning of the second administration: "Compared to this morning, would you say you feel better, worse, or the same?" with three predefined answers ("Better", "Worse", "The same").

The intraday test–retest reliability was assessed in the subgroup of adolescents who reported feeling the same on the two occasions only, using intra-class correlation coefficients (ICC), with a desired value > 0.70 [44]. While the EMOWI was designed to evaluate fluctuations in well-being in an EMA context, we restricted the sample for this analysis to optimize the evaluation of test–retest invariability.

Validity

Concurrent validity The Warwick-Edinburgh Mental Well-being Scale (WEMWBS) was selected for comparison because it has good psychometric properties [17], has been validated in adolescents [5], has been translated in French [45], and explores both hedonic and eudemonic aspects of well-being. Correlations between EMOWI and WEMWBS scores were calculated using Spearman's rank correlation coefficients, and a positive correlation was expected despite the different timeframes, as both scales measure similar concepts.

Correlations were checked between the EMOWI score and the scores of the Hospital Anxiety and Depression Scale (HADS) [46] and the Adolescent Depression Rating Scale

(ADRS) [47]. Both HADS and ADRS have been validated in French adolescent populations, and measure depressive symptoms, which are considered as opposite to mental well-being [19]. Negative correlations were expected.

Content validity Distribution of responses was examined to look for any floor (or ceiling) effect, defined as a clustering of more than 15% of participants at the lowest (or highest) score [44].

All quantitative analyses were conducted using R statistical software v4.0.3 for both the pictorial and the verbal versions of the scale, in order to compare their respective properties [23].

Ethical considerations

Prior to their participation, adolescents and their legal representatives were informed in writing about the study and their right to refuse to participate or to ask for the removal of their data at any time. Data from cognitive interviews that were considered as potentially identifying were not transcribed, to ensure complete anonymity.

The study was registered in accordance with French regulation following reference methodology MR004, in coordination with the Université de Lorraine Data Protection Office (no. 2021-156).

Results

Results are described according to the previously defined framework (Fig. 1). Table 1 presents an overview of the scale development qualitative process.

Phase 1: item development

Identification of domain(s) and item generation

The review identified 22 publications of interest, of which 6 reported the development of a scale, and 16 were conceptual papers. Six were specific to child or adolescent populations.

Analysis of these publications helped identify 16 major concepts related to mental well-being: competence, self-esteem, engagement, optimism, emotions, emotional stability, positive relationships, resilience, vitality, life or domain satisfaction, autonomy, environmental mastery, personal growth, purpose in life, health, and safety and supportive environment (Appendix).

The selection process led by the authors helped retain six dimensions covering both hedonic and eudemonic dimensions of adolescent mental well-being, and suitable to EMA: emotions, feeling of safety, positive relationships, vitality, self-esteem, and autonomy. A first set of eight items was drafted, to provide a momentary measurement of each dimension, with the dimension "emotions" being split into three items covering happiness, stress, and boredom.

Content validity assessment

All dimensions were endorsed by the experts. However, evaluation of relationships was perceived as insufficient by the adolescent expert, who suggested to add another item assessing the feeling of being valued. This referred to the concept of "connectedness", defined as the feeling of being accepted and loved by others in the work of Ryan, Huta and Deci [48] and in the EPOCH Measure of Adolescent well-being [9], and to an item of WEMWMS ("I've been feeling loved") [17], and was therefore unanimously approved by the authors. The item measuring "happiness" was reworded and

Table 1 Overview of the qualitative process carried out during the EMOWI development

Investigated scale version	Verbal version	Icons only	Pictorial version
Number of participants	5	8	7
Participants' characteristics	• 4 professionals • 1 adolescent	8 adolescents	7 adolescents
Type of data collection	Written electronic questionnaire	Cognitive interviews (think aloud)	Cognitive interviews (verbal probing)
Number of identified problems	5	2	1
Types of identified problems (according to Tourangeau's theory)	• Temporal comprehension (1) • Decision (1) • Response (2) • General comprehension (1)		• Temporal comprehension
Number of adjustments made in text	4	_	1
Number of adjustments made in icons	_	2	0

moved to the penultimate place to avoid confirmatory bias because of its understanding by two experts as a superior dimension.

Intercoder agreement was excellent, with Kappa scores greater than 0.80 for all cognitive processes (0.811 for general comprehension, 0.993 for temporal comprehension, 0.819 for decision, and 0.926 for response).

Phase 2: scale development

Development of icons

Icons were created for each anchor. For four items, two propositions were retained by the authors (Appendix).

Pre-testing of questions

Icons pre-test Icons for four items (self-esteem, loneliness, happiness, and feeling of being valued) were spontaneously understood by all adolescents. For three items with two available propositions (stress, autonomy, and safety), cognitive interviews helped select the most relevant ones, based on majority opinion (at least five out of eight adolescents). For two items, icons were improved after the first four interviews, thanks to suggestions made by adolescents: batteries were added to the icons representing energy, and the boredom icon look was emphasized. These modifications improved understanding during the next 3 interviews. Icons were then refined by graphic designers to ensure a harmonious presentation.

Item pre-test Previous iterative steps proved to be useful since questions and icons were well understood by all participants. Adolescents found the icons and slider helpful and enjoyable, in comparison with the verbal version *("I think icons are more interesting than words, because we don't need to think a lot and it's easier to relate with")*. They enjoyed the predefined steps on the slider, and agreed on the number of response options, supporting the use of a discrete scale *("I think it [the predefined steps] is a good idea, because if there are no steps, it is more difficult to place the slider, we don't know where to put it.")*. For all items, interviewed adolescents agreed on the intraday variability of the concepts and the importance of a momentary assessment: *"I think it [the feeling of safety] can change a lot throughout the day because it depends on many things: the people we are with, the kind of day we had, where we are …"*. Only one problem was evidenced, regarding the temporal comprehension of the item exploring "boredom", as one adolescent observed he couldn't be bored "at this moment", because he would be busy answering the questions. This was addressed by adding a temporal indication as an introduction to the scale, to refer to the last few minutes before completing the

questionnaire. All participants found the scale exhaustive regarding well-being.

Intercoder agreement was again excellent, with Kappa scores greater than 0.90 for all processes (0.903 for general comprehension; 0.941 for pictorial comprehension; 0.951 for temporal comprehension; 0.997 for decision; 0.905 for response).

Phase 3: scale evaluation

Sampling and survey administration

On the first administration, 387 and 383 adolescents completed the pictorial and the verbal versions of the EMOWI, respectively. Sample characteristics are described in Table 2.

Tests of dimensionality

Using the training dataset, CFA on the complete scale showed unsatisfactory fit (CFI: 0.903, TLI: 0.870, RMSEA: 0.081). Standardized loadings for all items were above 0.5, except for the item exploring boredom (standardized loading: 0.181), which was therefore removed. CFA estimations for the new 8-item scale were consistent with the unidimensional hypothesis, with CFI and TLI being above their desired values and RMSEA being under 0.08 for all datasets (training set, test set, and whole sample), for both the pictorial and the verbal versions (Table 3).

Reliability

Internal consistency Cronbach's alpha was 0.85 for the pictorial version and 0.86 for the verbal version, showing high internal consistency.

Table 2 Characteristics of the sample used for quantitative evaluation of the EMOWI

Characteristic	$N = 387$
Age (years)	
12	4 (1.0%)
13	192 (49.6%)
14	183 (47.3%)
15	8 (2.1%)
Gender	
Male	203 (52.5%)
Female	179 (46.3%)
Other	5 (1.3%)
School type	
General	380 (98.2%)
Adapted	5 (1.3%)
Other	2 (0.5%)

Table 3 Psychometric property estimations for the EMOWI pictorial and verbal versions

Scale version	CFI	TLI	RMSEA	Cronbach's alpha	ICC	Correlation with WEMWBS	Correlation with ADRS	Correlation with HADS	Correlation with feeling of being pressured
Pictorial	0.937	0.912	0.072	0.85	0.83	0.67[a]	−0.55[a]	−0.47[a]	−0.47[a]
Verbal	0.949	0.928	0.070	0.86	0.72	0.68[a]	−0.58[a]	−0.49[a]	−0.55[a]

CFI Comparative fit index, *TLI* Tucker–Lewis index, *RMSEA* Root mean square error of approximation, *ICC* Intra-class correlation coefficient, *WEMWBS* Warwick-Edinburgh Mental Well-being Scale, *ADRS* Adolescent Depression Rating Scale, *HADS* Hospital Anxiety and Depression Scale

[a]Statistically significant correlation ($p < 0.05$)

Test–retest reliability 376 adolescents completed the scale on the second administration. Due to logistic constraints, 138 participants completed the two measurement points in a row in the morning and were therefore excluded from the test–retest reliability analysis to avoid contamination by memorization of responses.

Among the remaining 238 adolescents, 136 (57.1%) reported to be feeling the same on the two occasions. ICC in this subgroup showed high reliability (Table 3).

Validity

Concurrent validity As hypothesized, a positive and significant correlation was evidenced with WEMWMS score, and negative and significant correlations were observed between the EMOWI global score, and scores obtained with ADRS and HADS (Table 3).

Content validity While responses for all items were skewed toward 10 after reverse scoring, this was less obvious on the global score, for which no ceiling effect was evidenced

(Fig. 2). No clustering > 15% was evidenced at 5, the slider departure position.

Comparison between the pictorial and verbal versions (Table 3, Appendix)

Both versions had similar psychometric properties.

Discussion

This paper followed the development process of the Ecological MOmentary Well-Being Instrument (EMOWI) for adolescents, the first scale of adolescents' mental well-being suitable for EMA. Given the current lack of appropriate measurement tools, the EMOWI could help better understand the impact of adolescents' emotional reactivity on health behaviors, as well as the effectiveness of dedicated mental health interventions [49].

The EMOWI showed proficient face validity, internal consistency, and reliability, after one item was removed. Correlations with existing instruments were consistent
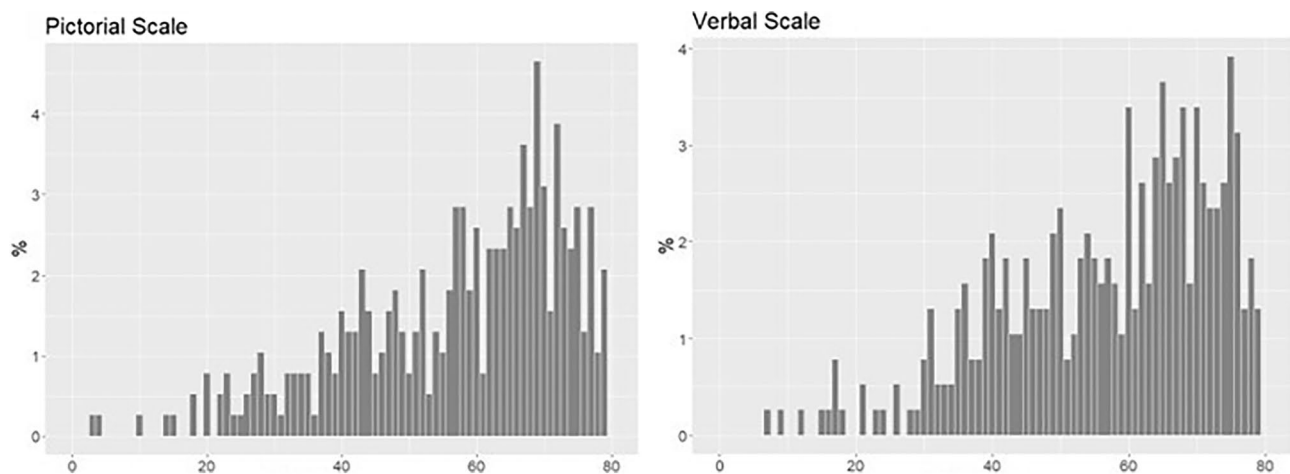

Fig. 2 Global score distribution for pictorial and verbal versions of EMOWI (median score for pictorial version: 62.0 [IQR: 47.0–70.5]; verbal version: 61.0 [47.0–71.0])

with predefined hypotheses. The fact that WEMWS and EMOWI investigate similar dimensions but use different timeframes and that the feeling of safety is not investigated by WEMWMS, may explain the positive, yet partial correlation < 0.70.

The pictorial version not only showed similar psychometric properties to the verbal version, but was also unanimously preferred by interviewed adolescents, as they found it easier to understand and relate with. Given the alleged increase in respondents motivation induced by pictorial scales, as well as their convenience in momentary measurement [23], these results further support the use of such scales in adolescent populations. Icons could also facilitate the adaptation of the EMOWI to other languages and cultural contexts, and help tackle lower participation rates in populations with low literacy.

Strengths

The EMOWI development was based on an innovative process combining recent recommendations by Boateng et al. [27] and Sauer et al. [23]. It included a comprehensive literature review, cognitive interviews, and an evaluation of the scale psychometric properties, which together helped progressively refine the instrument, until it seemed sufficiently understandable and relevant to adolescents. Qualitative findings also helped better understand quantitative results. For instance, the poor loading of the item assessing boredom was consistent with observations made by two adolescents during the pre-testing phase who mentioned that "boredom isn't necessarily a bad thing" and may therefore not be predictive of adolescents' momentary well-being, contrary to the initial assumption. The ceiling effect observed on item scores was also consistent with a comment from an expert, who stated that "during adolescence […], we tend to move towards extreme responses." Besides, no ceiling effect was evidenced on the global score, showing the EMOWI as a whole is well calibrated to detect different levels of well-being.

The EMOWI covers affective (happiness, stress, energy), cognitive (self-esteem, autonomy), as well as social aspects (loneliness, feeling of being valued, safety, autonomy) of adolescents' mental well-being. The development team included researchers from France and Canada, and adolescents from both countries participated to minimize the risk of cultural bias, for both text and icons. Finally, the sample used for quantitative evaluation included adolescents from heterogeneous socioeconomical and academic backgrounds, favoring understandability in various contexts.

To our knowledge, this report is also one of the firsts to provide a comparison of the psychometric properties of the pictorial and verbal versions of the same scale, showing

similar values and thus further supporting the use of pictorial scales, when appropriate [23].

Limitations

The EMOWI was not evaluated in a true EMA context in this study. Considering the consequent resources needed for EMA, as well as its more intrusive nature, we first aimed to ensure our scale had satisfactory basic properties. Besides, the EMOWI was specifically designed for EMA, and its comprehension as a momentary measure by adolescents was confirmed during qualitative interviews. These encouraging results are likely to support the suitability of the EMOWI for the detection of fluctuations in adolescents' well-being using EMA.

Further research is also needed regarding the most effective layout for sliders used in self-measurement tools. For instance, adding visual cues along the slide steps, as in the Affective Slider [22], might help respondents better convey their feeling on the scale and reduce floor or ceiling effects. Besides, despite a transcultural development, quantitative evaluation of the EMOWI could not be conducted in Canada due to restrictions related to the Covid-19 pandemic. However, since the EMOWI psychometric properties in the French sample were satisfactory, it is likely future testing in Canada will show similar results.

Conclusions

To this day, the EMOWI is the first instrument to provide a measurement of momentary well-being in adolescents. Its innovative nature, strong psychometric performance, shortness, and easiness of use make it an excellent candidate for use in EMA. Future work should test the instrument in EMA, to further confirm its potential to investigate health behaviors and mental health in adolescents.

commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability Data, models, and methodologies used in this research may be available from the corresponding author upon reasonable request.

Declarations

Conflict of interest YK holds shares in Polygon Research Inc., the company that markets the Research Suite application. The other authors have no relevant financial or non-financial interests to disclose.

References

1. World Health Organization. Mental health: strengthening our response. https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response (accessed June 22, 2021).

2. Arbour-Nicitopoulos, K. P., Faulkner, G. E., & Irving, H. M. (2012). Multiple health-risk behaviour and psychological distress in adolescence. *Journal of Canadian Academy of Child and Adolescent Psychiatry, 21*, 171–178.

3. Griffin, K. W., Scheier, L. M., Botvin, G. J., & Diaz, T. (2001). Protective role of personal competence skills in adolescent substance use: Psychological well-being as a mediating factor. *Psychology of Addictive Behaviors, 15*, 194–203.

4. Shedler, J., & Block, J. (1990). Adolescent drug use and psychological health. A longitudinal inquiry. *American Psychology, 45*, 612–630. https://doi.org/10.1037//0003-066x.45.5.612

5. Clarke, A., Friede, T., Putz, R., Ashdown, J., Martin, S., Blake, A., et al. (2011). Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Validated for teenage school students in England and Scotland. A mixed methods assessment. *BMC Public Health, 11*, 487. https://doi.org/10.1186/1471-2458-11-487

6. Tomyn, A. J., Stokes, M. A., Cummins, R. A., & Dias, P. C. (2020). A Rasch analysis of the personal well-being index in school children. *Evaluation and the Health Professions, 43*, 110–119. https://doi.org/10.1177/0163278718819219

7. Ravens-Sieberer, U., Erhart, M., Rajmil, L., Herdman, M., Auquier, P., Bruil, J., et al. (2010). Reliability, construct and criterion validity of the KIDSCREEN-10 score: A short measure for children and adolescents' well-being and health-related quality of life. *Quality of Life Research, 19*, 1487–1500. https://doi.org/10.1007/s11136-010-9706-5

8. Viejo, C., Gómez-López, M., & Ortega-Ruiz, R. (2018). Adolescents' psychological well-being: A multidimensional measure. *IJERPH, 15*, 2325. https://doi.org/10.3390/ijerph15102325

9. Kern, M. L., Benson, L., Steinberg, E. A., & Steinberg, L. (2016). The EPOCH measure of adolescent well-being. *Psychological Assessment, 28*, 586–597. https://doi.org/10.1037/pas0000201

10. Ahmed, S. P., Bittencourt-Hewitt, A., & Sebastian, C. L. (2015). Neurocognitive bases of emotion regulation development in adolescence. *Developmental Cognitive Neuroscience, 15*, 11–25. https://doi.org/10.1016/j.dcn.2015.07.006

11. Young, K. S., Sandman, C. F., & Craske, M. G. (2019). Positive and negative emotion regulation in adolescence: Links to anxiety and depression. *Brain Sci.* https://doi.org/10.3390/brainsci9040076

12. Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*, 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

13. van Genugten, C. R., Schuurmans, J., Lamers, F., Riese, H., Penninx, B. W., Schoevers, R. A., et al. (2020). Experienced burden of and adherence to smartphone-based ecological momentary assessment in persons with affective disorders. *J Clin Med.* https://doi.org/10.3390/jcm9020322

14. Beute, F., & de Kort, Y. A. W. (2018). The natural context of wellbeing: Ecological momentary assessment of the influence of nature and daylight on affect and stress for individuals with depression levels varying from none to clinical. *Health & Place, 49*, 7–18. https://doi.org/10.1016/j.healthplace.2017.11.005

15. Maher, J. P., Doerksen, S. E., Elavsky, S., & Conroy, D. E. (2014). Daily satisfaction with life is regulated by both physical activity and sedentary behavior. *Journal of Sport & Exercise Psychology, 36*, 166–178. https://doi.org/10.1123/jsep.2013-0185

16. de Vries, L. P., Baselmans, B. M. L., & Bartels, M. (2021). Smartphone-based ecological momentary assessment of well-being: A systematic review and recommendations for future studies. *Journal of Happiness Studies, 22*, 2361–2408. https://doi.org/10.1007/s10902-020-00324-7

17. Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., et al. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes, 5*, 63. https://doi.org/10.1186/1477-7525-5-63

18. Ross, D. A., Hinton, R., Melles-Brewer, M., Engel, D., Zeck, W., Fagan, L., et al. (2020). Adolescent well-being: A definition and conceptual framework. *Journal of Adolescent Health, 67*, 472–476. https://doi.org/10.1016/j.jadohealth.2020.06.042

19. Huppert, F. A., & So, T. T. C. (2013). Flourishing across Europe: Application of a new conceptual framework for defining well-being. *Social Indicators Research, 110*, 837–861. https://doi.org/10.1007/s11205-011-9966-7

20. Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior, 43*, 207–222.

21. Remschmidt, H. (1994). Psychosocial milestones in normal puberty and adolescence. *Hormone Research, 41*(Suppl 2), 19–29. https://doi.org/10.1159/000183955

22. Betella, A., & Verschure, P. F. M. J. (2016). The affective slider: A digital self-assessment scale for the measurement of human emotions. *PLoS One, 11*, e0148037. https://doi.org/10.1371/journal.pone.0148037

23. Sauer, J., Baumgartner, J., Frei, N., & Sonderegger, A. (2021). Pictorial scales in research and practice: A review. *European Psychologist, 26*, 112–130. https://doi.org/10.1027/1016-9040/a000405

24. Kunin, T. (1955). The construction of a new type of attitude measure. *Personnel Psychology, 8*, 65–77. https://doi.org/10.1111/j.1744-6570.1955.tb01189.x

25. Bennett, K. B., & Walters, B. (2001). Configural display design techniques considered at multiple levels of evaluation. *Human Factors, 43*, 415–434. https://doi.org/10.1518/001872001775898304

26. Baumgartner, J., Sonderegger, A., & Sauer, J. (2019). No need to read: Developing a pictorial single-item scale for measuring perceived usability. *International Journal of Human-Computer Studies, 122*, 78–89. https://doi.org/10.1016/j.ijhcs.2018.08.008

27. Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for

developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, 149. https://doi.org/10.3389/fpubh.2018.00149

28. van Roekel, E., Keijsers, L., & Chung, J. M. (2019). A review of current ambulatory assessment studies in adolescent samples and practical recommendations. *Journal of Research on Adolescence, 29*, 560–577. https://doi.org/10.1111/jora.12471

29. Diener, E., Suh, E., Lucas, R., & Smith, H. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*, 276–302. https://doi.org/10.1037/0033-2909.125.2.276

30. Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology, 52*, 141–166. https://doi.org/10.1146/annurev.psych.52.1.141

31. Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., et al. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research, 97*, 143–156. https://doi.org/10.1007/s11205-009-9493-y

32. Ryff, C. D., & Keyes, C. L. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology, 69*, 719–727. https://doi.org/10.1037//0022-3514.69.4.719

33. Courtwright, S. E., Makic, M. B. F., & Jones, J. (2020). Emotional wellbeing in youth: A concept analysis. *Nursing Forum, 55*, 106–117. https://doi.org/10.1111/nuf.12404

34. Keyes, C. L. M. (2006). The subjective well-being of america's youth: toward a comprehensive assessment. *Adolescent & Family Health, 4*, 3–11.

35. Rebar L. (n.d). LibGuides: Nursing - Literature reviews & systematic reviews: backward & forward reference searching. https://libguides.fau.edu/c.php?g=966176&p=7013756. Accessed 23 July 2021.

36. National Research Council. (1984). Cognitive aspects of survey methodology: Building a bridge between disciplines. The National Academies Press. https://doi.org/10.17226/930.

37. Willis, G. (2021). *Cognitive interviewing*. Thousand Oaks.

38. Murphy, M., Hollinghurst, S., & Salisbury, C. (2018). Qualitative assessment of the primary care outcomes questionnaire: A cognitive interview study. *BMC Health Services Research, 18*, 79. https://doi.org/10.1186/s12913-018-2867-6

39. McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochem Med (Zagreb), 22*, 276–282.

40. Patrick, D. L., Burke, L. B., Gwaltney, C. J., Leidy, N. K., Martin, M. L., Molsen, E., et al. (2011). Content Validity—Establishing and Reporting the Evidence in Newly Developed Patient-Reported Outcomes (PRO) Instruments for Medical Product Evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 2—Assessing Respondent Understanding. *Value in Health, 14*, 978–988. https://doi.org/10.1016/j.jval.2011.06.013

41. 2.4.1 Lemmatisation — IRaMuTeQ. (n.d). http://www.iramuteq.org/documentation/html/copy4_of_1-2-principe-de-fonctionnement-du-logiciel. Accessed 23 July 2021.

42. Willis, G. B., & Artino, A. R. (2013). What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. *Journal of Graduate Medical Education, 5*, 353–356. https://doi.org/10.4300/JGME-D-13-00154.1

43. Mokkink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, de Vet HC, et al. (n.d). COSMIN Study Design checklist for Patient-reported outcome measurement instruments, *32*.

44. Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60*, 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

45. Trousselard, M., Steiler, D., Dutheil, F., Claverie, D., Canini, F., Fenouillet, F., et al. (2016). Validation of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) in French psychiatric and general populations. *Psychiatry Research, 245*, 282–290. https://doi.org/10.1016/j.psychres.2016.08.050

46. White, D., Leach, C., Sims, R., Atkinson, M., & Cottrell, D. (1999). Validation of the hospital anxiety and depression scale for use with adolescents. *British Journal of Psychiatry, 175*, 452–454. https://doi.org/10.1192/bjp.175.5.452

47. Revah-Levy, A., Birmaher, B., Gasquet, I., & Falissard, B. (2007). The Adolescent Depression Rating Scale (ADRS): A validation study. *BMC Psychiatry, 7*, 2. https://doi.org/10.1186/1471-244X-7-2

48. Ryan, R., Huta, V., & Deci, E. (2013). Living well: a self-determination theory perspective on Eudaimonia. The Exploration of Happiness, Happiness Studies Book Series, pp. 117–39. https://doi.org/10.1007/978-94-007-5702-8_7.

49. Schueller, S. M., Aguilera, A., & Mohr, D. C. (2017). Ecological momentary interventions for depression and anxiety. *Depression and Anxiety, 34*, 540–545. https://doi.org/10.1002/da.22649