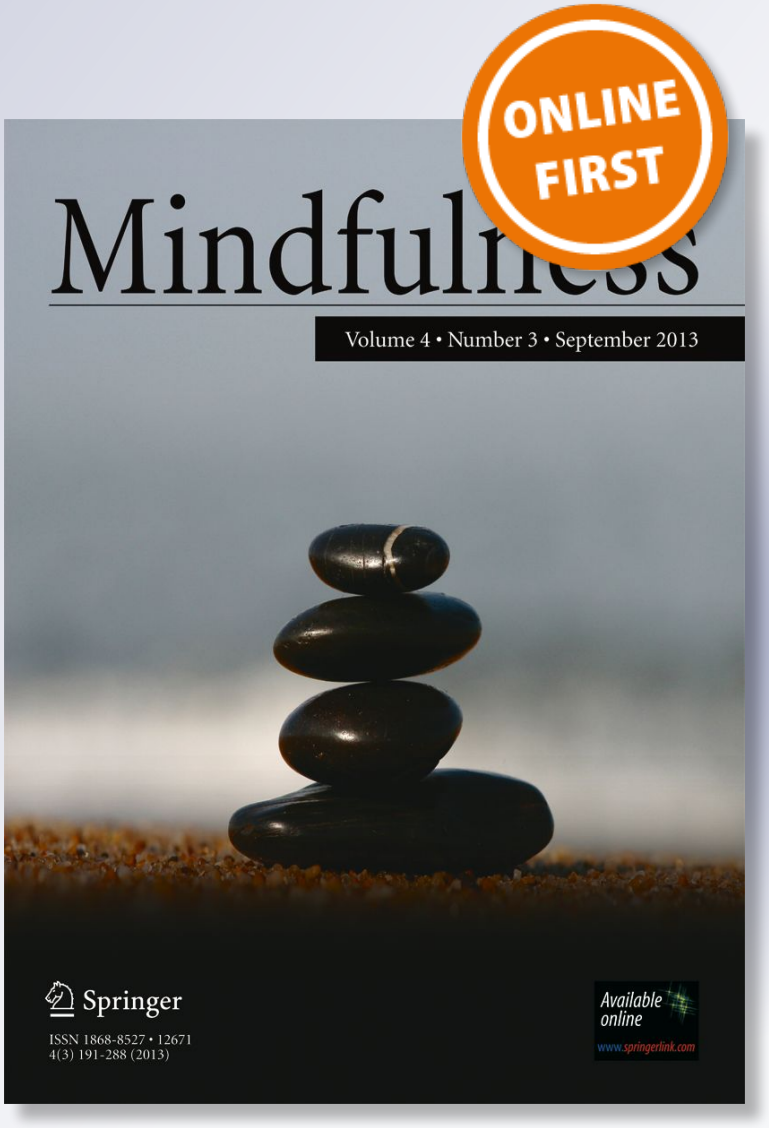
Validation of a Scale for Assessing Social Validity in Mindfulness-Based Educational Programs

Luis López-González, David Herrero-Fernández, Alberto Amutio, Txemi Santamaría & William Van Gordon

Mindfulness

ISSN 1868-8527

Mindfulness
DOI 10.1007/s12671-019-01138-2



Springer

Mindfulness
https://doi.org/10.1007/s12671-019-01138-2

ORIGINAL PAPER

Validation of a Scale for Assessing Social Validity in Mindfulness-Based Educational Programs

Luis López-González[1] · David Herrero-Fernández[2] · Alberto Amutio[3] · Txemi Santamaría[4] · William Van Gordon[5]

Abstract

Objectives Social validity (SV) is a concept used in intervention research and is concerned with the overall acceptability, relevance, and utility of an intervention to all intervention stakeholders. SV not only takes into account efficacy in respect of the pre-defined study outcomes, but also participants' perceptions of the intervention as well as the wider social context in which it will be applied. There are a growing number of mindfulness-based educational programs (MBEPs) being empirically evaluated and implemented in educational settings. However, due to a lack of scientifically validated instruments that can assess SV in MBEPs, a systematic evaluation of SV in such programs has not been undertaken to date. The aim of this study was to investigate the psychometric properties of the *Social Validity Scale of Mindfulness-Based Programs for Adolescents* (*Escala de Validez Social de Programas de Mindfulness para Adolescentes—EVSPM-A*), composed of 20 items.

Methods The sample comprised 512 compulsory secondary education and high school students (mean age = 14.5; *SD* = 1.57) from three Spanish educational centers that had completed an MBEP known as the *TREVA Program*.

Results The final version of the scale showed good psychometric properties and factor analyses yielded five factors: *global impact-satisfaction*, *acceptance and viability*, *individual perceived effectiveness*, *perceived classroom climate; training feasibility*, and *applicability of techniques*.

Conclusions The EVSPM-A appears to be a suitable means of assessing SV in MBEPs delivered to adolescents. Using the EVSPM-A to evaluate SV can help improve the implementation and long-term efficacy of MBEPs.

Keywords Assessment · Social validity · Mindfulness-based interventions · Educational programs

Mindfulness is reported to be beneficial for adolescents in clinical (Tan 2016), health promotion (Greenberg and Harris 2012), social (Beauchemin et al. 2008), substance abuse (Himelstein 2016), judicial (Leonard et al. 2013), and human performance contexts (Baltzell 2016). Mindfulness is also reported to have applications in adolescent educational contexts (Wisner 2017) where it can promote healthy brain functioning (Davidson et al. 2012), heighten attentional skills (Sanger and Dorjee 2015), and improve executive functioning (Flook et al. 2010). Accordingly, mindfulness-based interventions (MBIs) and mindfulness-based education programs (MBEPs) are increasingly being employed in adolescent educational contexts (Shonin et al. 2012). In fact, there is estimated to be at least 30 different mindfulness programs for children and adolescents around the world (Ager et al. 2015). In the UK, nearly 50% of school-age children engage in some form of mindful activity, and approximately 50% of secondary school teachers use mindfulness with children in their classes (Stone 2014). This is in line with a global trend reflecting an increase of interest into the applications of mindfulness in school-based settings (Sapthiang et al. 2019; Stone 2014).

Although interest into the applications of MBIs for adolescent learners is growing, several meta-analytical studies indicate that there are major shortcomings that undermine research outcomes in this respect (Kallapiran et al. 2015; Klingbeil et al. 2017a, b; Zenner et al. 2014). Firstly, most studies of

✉ Alberto Amutio
alberto.amutio@ehu.eus

[1] Institut de Desenvolupament Professional, University of Barcelona, Barcelona, Spain

[2] Universidad Europea del Atlántico, Santander, Spain

[3] Facultad de Relaciones Laborales y Trabajo Social, University of the Basque Country (UPV/EHU), 48940 Leioa, Bizkaia, Spain

[4] Faculty of Psychology and Education, University of Deusto, Bilbao, Spain

[5] Human Sciences Research Centre, University of Derby, Derby, UK

Published online: 30 March 2019

this nature focus on the effects of isolated MBIs which means that it is difficult to pool outcomes due to different models of mindfulness being taught (Pallozzi et al. 2016). There are also issues relating to the reliability, validity, and general suitability of many of the measures employed in MBI adolescent education-focused studies (Bluth and Blanton 2015; Shonani and Slone 2013). Furthermore, very few studies have sought to assess the impact of mindfulness on aspects of the learning environment including, for example, classroom climate, learning culture, bullying, and academic performance (Dariotis et al. 2016). An additional issue relates to the pedagogic quality of MBIs that is often not evaluated in terms of the suitability of a given MBI (or MBI instructors' teaching style) for a particular population of adolescent learners (Sainio et al. 2018; Semple et al. 2017).

Consequently, before delivering MBIs/MBEPs to adolescent learners, it would be prudent to comprehensively assess the suitability, effectiveness, risks, and limitations of these programs (Eklund et al. 2017; López-González et al. 2018a). Indeed, educational interventions should ideally respond to a pre-intervention needs detection or analysis, which is practically non-existent in most education-focused MBIs (López-González et al. 2016a; Rennekamp and Arnold 2009; Schonert-Reichl and Roeser 2016). Conversely, MBEPs are typically designed specifically for children and adolescents and include an assessment of needs based on an educational setting (context). The MBEP is then designed to meet these needs (process), and the intervention is subsequently evaluated, including with reference to its wider social validity (product). Furthermore, the design, implementation, and evaluation of MBEPs is typically based on both quantitative and qualitative methodologies, as well as on an evaluation of the classroom climate, stakeholder satisfaction, instructor style, and whether users feel equipped to continue practicing mindfulness following completion of the intervention (López-González et al. 2016a; Worthen and Luiselli 2017). Given the unique design and evaluation considerations of MBEPs, they should arguably be considered a specific sub-group of MBIs. Indeed, relating to MBEPs in such a manner would likely increase cohesion across MBEPs as well as foster a better understanding of how to evaluate the wider utility of MBEPs in order to maximize benefit to young learners (Forman and Soloff 2014).

At present time, two important questions that remain unanswered are whether mindfulness-based educational programs have utility and, are they effective? Social validity (SV), which relates to whether the intervention is effective and has utility in real-world contexts (Carter 2009), is a means of holistically evaluating a given intervention in a specific setting. Rather than solely focusing on efficacy in respect of a small number of predefined outcomes, SV takes into account participants' perceptions of the intervention as well as the wider social context in which it will be applied (Kern and Manz 2004; Marchant et al. 2013). According to Wolf (1978), social validity has three essential facets: utility, social appropriation of results, and relevance. In addition to using psychometric scales (Carroll and St. Peter 2014) and qualitative techniques, SV is also measured by evaluating the maintenance of the program through time as well as its acceptability among direct and indirect stakeholders (Hanley 2012; Jennings et al. 2017).

Aspects of SV have been used to guide the development and assessment of several MBEPs delivered in elementary schools (Shababi 2014), special education centers (Hokke et al. 2014), and sports performance contexts (Baltzell 2016). More specifically, feasibility, acceptability, and effectiveness were evaluated in a *Stress Reduction and Mindfulness Curriculum* as well as in a *Mindful Moment* initiative implemented in an elementary school in Baltimore (Mendelson et al. 2010). Examples of additional mindfulness-informed programs in which different aspects of SV have been evaluated include the (i) *Master Mind and Moment Program* (Parker and Kupersmidt 2016) where feasibility, acceptability, and low cost were assessed; (ii) *Mindfulness and Mind-Body Skills for Children* (MMBS) (Sheinman et al. 2011) where feasibility and acceptability were measured; (iii) *Mindful Schools* (Liehr and Diaz 2010) where acceptability was assessed; (iv) *Still Quiet Place* were feasibility was assessed (Saltzman and Goldin 2008); (v) *Inner Explorer* program where viability was evaluated; (vi) *Wellness and Resilience Program* where effectiveness and sustainability were assessed (Semple et al. 2017); and (vii) Spanish-based *TREVA Program* where impact, effectiveness, acceptability, feasibility, sustainability, and transportability were measured (López González 2010; López-González et al. 2016a).

Findings from studies evaluating the aforementioned TREVA Program (López-González et al. 2016a) support the view that teachers themselves should teach mindfulness to students, thus confirming the need to include an evaluation of teacher effectiveness and training as part of MBEP implementation (Emerson et al. 2017; Zenner et al. 2014). The TREVA is based on nine different taught components, comprising seven relaxation, meditation and mindfulness skills (attention, letting go [relaxation], breathing, visualization, sensory awareness, postural awareness and energy awareness), and two special competencies (emotional management, and mindfulness [including body awareness]). The program trains teachers for a total of 20 h over the course of 10–12 weeks, plus an additional 10 h of online training. The first phase of the TREVA program is devoted exclusively to the learning of relaxation-mindfulness (REMIND) techniques as well as their application to the conscious management of the classroom (López-González et al. 2018a). The second phase focuses on how students can learn and apply the *REMIND* techniques. A third phase, called *Action Plan*, specifically focuses on how to tailor the mindfulness program to the student population in question, which is in line with evaluating SV and the wider utility of MBEPs. The fact that the teachers

themselves (i.e., rather than external experts) oversee students' training, encourages more teachers and pedagogy stakeholders to become involved in program delivery. The TREVA program represents one of the few attempts to integrate mindfulness practice into the educational curricula rather than offering it as an isolated component.

Although research in the MBEP field has been expanding over the last two decades, the empirical evaluation of SV in MBEPs remains underdeveloped. Furthermore, only a limited number of MBI studies have evaluated specific components of SV (e.g., Kuyken et al. 2013), and a purpose-designed SV assessment instrument for MBIs or MBEPs is yet to be developed. Indeed, most studies have included an assessment of SV as a non-primary outcome and often via the inclusion of just one or two SV-related items (Luiselli et al. 2017; Worthen and Luiselli 2017). Furthermore, some of these studies used unvalidated questionnaires that were designed by the corresponding study authors, which meant that in some cases, the psychometric properties of the scale were not known (Metz et al. 2013). In other cases, studies have used generic SV questionnaires that have not purposely sought to assess the SV construct in the specific context of mindfulness practice (Felver et al. 2016). A means of rigorously assessing SV in the context of MBEPs is necessary not only to enhance intervention design, effectiveness, and acceptability to participants, but because perceptual changes and reappraisal processes elicited during mindfulness can change how participants relate to and evaluate a given experience or intervention. Furthermore, these perceptual changes triggered by mindfulness are understood to be different between adults and individuals transitioning through the developmentally demanding period of adolescence (Sapthiang et al. 2019). Therefore, a measure of SV that is sensitive to the evaluative processes of mindfulness practitioners, and in particular adolescent mindfulness practitioners, would lead to a more accurate understanding and appraisal of SV in the context of MBEPs.

Consequently, the purpose of this study was to assess the psychometric properties of the *Social Validity Scale of Mindfulness-Based Programs for Adolescents* (*Escala de Validez Social de Programas de Mindfulness para Adolescentes; EVSPM-A*). The instrument is the first SV tool that is specifically designed for Spanish adolescents (in early and late adolescence) receiving mindfulness training.

Method

Participants

The sample comprised 512 students (male = 54.5%, female = 45.5%) aged between 12 and 18 years old ($M = 14.54$; $DT = 1.57$). A total of 10.5% of participants were completing the first year of High School/Baccalaureate and the remainder of participants—all of whom were aged between 12 and 16 years—were completing Compulsory Secondary Education (first year = 20.1%, second year = 25.2%, third year = 21.3%, fourth year = 22.9%). Participants were recruited from three different Spanish education centers (two public institutes of Barcelona and one semi-public center in Bilbao), which had taken part in the TREVA Program. Participation was on a voluntarily basis and all participants provided informed consent.

Procedure

The final scale, which sought to assess SV of intervention procedure (i.e., in terms of users´ perceptions of design and development, as well as acceptability of the intervention), was the product of a set of preliminary steps. After an exhaustive review of the SV literature, SV terms most frequently used by researchers were selected. Key SV terms were deemed to be *effectiveness*, *impact*, *satisfaction*, *election*, *choice*, *maintenance*, *importance*, *completeness*, *acceptability*, *viability*, *feasibility*, *and sustainability*. The following six SV terms were then short-listed based on their relevance to the implementation of MBIs and MBEPs: *effectiveness*, *impact*, *viability*, *feasibility*, *acceptance*, and *applicability*. The following criteria were subsequently applied during the course of integrating these terms into scale items: (a) items must explicitly refer to SV key features of MBIs/MBEPs and (b) items must reflect users' perceptions of design and development (i.e., and not just the program effects).

The original scale was composed of 22 items and was developed during the course of a Doctoral Dissertation (psychometric properties were not investigated at that point; López-González 2010). All items were independently reviewed by three mindfulness experts. Next, a pilot study was conducted on a sample of students ($n = 35$) to rule out items that duplicated assessment components or that were too difficult for participants to understand. Two items were eliminated following the pilot study, and a total of 20 items were included in the revised version of the scale.

Measures

The present study was conducted following completion of three different intervention delivery iterations of an MBEP known as the TREVA Program (López-González et al. 2016a). In order to validate the scale, various instruments that are administered during the standard delivery of the TREVA Program (i.e., as a means of evaluating its effectiveness in school and adolescent settings) were administered in conjunction with the EVSPM-A.

Social Validity Scale of Mindfulness-Based Programs for Adolescents (*Escala de Validez Social de Programas de Mindfulness para Adolescentes—EVSPM-A*). The 20-item

EVSPM-A measures five factors of SV in MBEPs and MBIs programs that have been adapted to educational contexts. The scale utilizes four possible answers (1 = nothing, 2 = a little, 3 = quite a lot, 4 = a lot) to measure the effectiveness of an intervention or program not only in terms of the individual effects, but also the impact on the entire classroom. To our knowledge, this is the first instrument that evaluates both individual and group social validity effects. The validity and reliability of the scale is the subject of this study and the development of the scale is described below.

Relaxation-Mindfulness Skills and states Questionnaire for Adolescents (*Escala de Habilidades y Estados de Relajación-Mindfulness para Adolescentes—EHERM-A*; López-González et al. 2016b). The EHERM-A features 19 items (e.g., "I have followed the exercise very carefully") that measure the focusing and passivity-receptivity dimensions of relaxation and mindfulness. The scale also evaluates REMIND states (e.g., "Happiness," "Disengagement," "Energy"). Items are scored against four possible responses (1 = never, 2 = sometimes 3 = frequently, 4 = always), and overall reliability is very good ($\alpha = .90$).

School Relaxation and Mindfulness Habits Brief Questionnaire (*Cuestionario de Hábitos de Relajación y Meditación Escolar—CBHRME*; López-González et al. 2016c). The scale comprises 11 items and evaluates habits relating to relaxation and mindfulness in terms of *personal habits* (e.g., "I look for peace calm and tranquility sensations"), *family habits* (e.g., "There is a peaceful and calm atmosphere at home"), and *school habits* (e.g., "We do some rest or relaxation exercises in class"). Items are scored on a 4-point Likert scale (1 = never, 2 = sometimes 3 = frequently, 4 = always). The psychometric properties and internal consistency ($\alpha = .72$) are good (López-González et al. 2016c).

Relaxation and Mindfulness Scale for Adolescents (*Escala de Relajación y Mindfulness para Adolescentes - EREMIND-A*) (López-González et al. 2018b). This 18-item Likert-type scale measures three different factors of relaxation and mindfulness for adolescents: (i) *attention and concentration in the present moment* (F1) (e.g., "In the classroom my mind is unoccupied"), (ii) *relaxation*(*F2*) (e.g., "I am scare of relaxing totally"), and (*iii*) *sensory awareness*, *contemplation/interiority* (F3) (e.g., "I like to pay attention to doing simple things and fully feeling them"). The scale demonstrates a reliability of $\alpha = .72$, and there are four possible responses (1 = nothing, 2 = a little, 3 = quite, 4 = a lot).

Classroom Climate Brief Scale (*Cuestionario Breve de Clima de Clase—CBCC*; López González and Bisquerra 2013). The *CBCC* is an 11-item Likert scale with four possible responses: *never*, *sometimes*, *often*, and *always*. Items are clustered into two dimensions of *group cohesion* (satisfaction, involvement, and cohesion among students) and *group leadership* (order and organization, task orientation, and the relationship between teacher and learners). Sample items include "There is a good atmosphere in class" and "There is the necessary silence for each subject"). The internal consistency of the scale is good ($\alpha = .83$).

Posture

Participants' level of attention to posture and its suitability for the task at hand (e.g., writing) was evaluated through a questionnaire designed for the purposes of this study. A template with five levels (1 = very incorrect; 2 = incorrect; 3 = not incorrect nor correct; 4 = correct; 5 = very correct.) was used. Following a process of discussion and standardization involving several of the present authors, the test was independently performed by direct observation by one member of the research team (i.e., who was not involved in facilitating the TREVA program). The observation was based on Da Fonceca's (1998) criteria and included the observation of two essential aspects: (i) optimal degree of muscular tone and (ii) global posture for the task (both feet on the floor, back straight without tension, arms well supported on the table while writing). The observer evaluated the students' posture to the extent that it met these criteria.

Closed-Eyes Test

This involved using a timer to measure the time students had their eyes shut while they undertook a 5-min relaxation exercise.

Disciplinary Sanctions

Information relating to disciplinary sanctions and expulsions was obtained from the official records of educational centers. This was obtained at the end of the program.

Data Analyses

Exploratory factor analysis (EFA) was conducted using a sample of 420 students. The EFA was carried out through a parallel analysis (PA) using the software Factor 9.2 (Lorenzo-Seva and Ferrando 2013). The method for factor extraction was the Robust Diagonally Weighted Least Squares, which suggested the existence of 5 factors that accounted for the 63.40% of the variance. In order to confirm the initial factorial structure of the EVSPM-A, a confirmatory factor analysis (CFA) was performed ($n = 512$). The CFA was conducted using Lisrel 8.80 software (Jöreskog and Sörbom 1997), and it followed the Robust Diagonally Weighted Least Squares method for estimating parameters. This is the most recommended method when multivariate normality is not guaranteed, as in the case with Likert-type items (Mîndrilă 2010). Following this, the results of the CFA were interpreted through the global fit of the model, considering the following six fit

indexes: (i) the quotient between $\chi 2$ and the degrees of freedom of the model, which should be lower than 3 to be considered a good fit (Carmines and McIver 1981); (ii) the root mean squared error of approximation (RMSEA); (iii) the standardized root mean square residual (SRMR), which should be close to 0.06 (Hu and Bentler 1999; Steiger 2007); (iv) the Non-Normed Fit Index (NNFI); (v) the Comparative Fit Index (CFI); and (vi) the Goodness of Fit Index (GFI), which should be above 0.95 (Hu and Bentler 1999).

Once a final factorial solution was established, factor loadings and error of estimation of the items were represented in a path diagram. Internal consistency (Cronbach's α), construct reliability (CR), and average variance explained (AVE) of the factors were also calculated. Descriptive statistics of both items and factors were also performed, including mean, standard deviation, skewness, and item-factor correlations.

Convergent criterion validity was calculated ($n = 420$) by correlating the mean of the scale with the means of the EREMIND-A, CBHRME, *Posture* observation test, and the CBCC. Divergent validity was obtained using a subsample of 335 students by correlating the mean of the scale with the number of disciplinary sanctions and penalties. In addition, discriminant validity of the five factors was calculated through the comparison of AVE with the shared variance between each pair of factors. If the root square of the AVE of *X* factor is higher than the correlation coefficient between *X* and *Y* factors, then the discriminant validity of the *X* factor is supported (Fornell and Larcker 1981).

Predictive validity was obtained 2 months later using a sample of 342 students by measuring how much the five factors and the global mean of the scale could predict the acquired level on REMIND skills using the EHERM-A. This was performed after a 15-min mindfulness exercise based on the TREVA Program (López-González et al. 2016a). Regression analyses were run to measure the predictive validity of each factor, after controlling for gender and academic level.

Furthermore, incremental validity was analyzed through two regression models by hierarchical approximation and analyzing zero, partial and semi-partial correlations between predictive variables and criterion. The level of abilities and states of REMIND measured with the EHERM-A were adopted as criterion. Gender and academic level were checked and the scale was compared with two other instruments measuring REMIND: the CBHRME and the EREMIND-A. Thus, in the first regression model, the variable EHERM-A was taken as the criteria after controlling for gender and academic level (step 1). Then, EREMIND-A and CBRME were introduced (step 2) followed by the mean EVSPM-A in step 3. In the second regression model, steps 1 and 2 above were followed, but in step 3, the five factors were introduced. If a statistical change of R^2 was observed through ANOVA, its practice utility was analyzed by interpreting the semi-partial correlation as the unique absolute improvement of this variable, where the minimum value should be between 0.15 and 0.20 (Husley and Meyer 2003).

Finally, differences due to gender and academic level were analyzed using a 2 (gender) × 5 (academic level) MANOVA. Then, both gender and academic level were taken as independent variables, whereas the individual factors of the EVSPM-A were included as dependent variables. Multivariate and univariate comparisons were conducted, and both significance and effect size (η^2) were analyzed. Cohen's criteria were employed to interpret effect sizes (where 0.01 to 0.04 = small effect, 0.05 to 0.14 = medium effect, greater than 0.14 = large effect).

Results

EFA yielded five factors that corresponded to the key SV terms most frequently used by researchers in the literature (López-González 2010): (i) *Global Impact-Satisfaction*, *Acceptance and Viability* (F1: items 1, 3, 7,9,13, 15, and 18; e.g., "I liked to do relaxation-mindfulness before starting some classes"); (ii) *Perceived Individual Effectiveness* (F2: items 2, 10, 12; e.g., "I had more energy in class after practicing relaxation and mindfulness"), which measures concentration and task-oriented energy; (iii) *Perceived Classroom Climate* (F3: items 4, 14 and 17; e.g., "After doing relaxation/mindfulness the atmosphere in class is more enjoyable."), which relates to aspects such as the learning environment and relationship with the teacher; (iv) *Training Feasibility* (F4: items 5, 11 and 19; e.g., "The time devoted to each practice was adequate"), which focuses on didactic aspects as well as the teaching style of the instructor (Semple et al. 2017; Jennings et al. 2017); and (v) *Applicability of Techniques* (F5: items 6, 8, 16, 20; e.g., "Techniques are easy to be learned").

For the CFA, three models were constructed and compared. The first was a five-factor model, according to the empirical result outlined above. The second was a one-factor model, given that the five factors were supposed to be correlated to each other. The third model was a bifactorial model, combining the previous two models. This option was included instead of a high-order model because in the bifactorial model, the second-order factor is made from the first-order factors, rather than from the individual items. Furthermore, the introduction of a second-order factor does not reflect usually relevant changes in goodness-of-fit indices. A bifactor model permits the generation of both a five-factor structure and a one-factor-structure at the same time, analyzing the specific amount of variance that each item loads in the first part (regarding the five factors) and the second part (one-factor) (Wiesner and Schanding 2013). In this last scenario, the five factors were conserved, and a sixth new factor was introduced. Each item loaded both on its respective factor and the new factor.

The five-factor solution showed a good fit to the data according to all of the goodness-of-fit indexes that were assessed

through CFA. Conversely, both the one-factor solution and the bifactorial solution did not fit satisfactorily. In the case of the one-factor solution, only the GFI met the cut-off. In the case of the bifactorial model, it showed an overfit to the data, as the value of the degrees of freedom of the model was higher than the chi-square value, and the RMSEA had a value of 0.000 (90% I.C. = 0.000–0.000) (Bollen and Curran 2005). Furthermore, this model showed very low (non-significant) factor loadings in the case of the new (i.e., sixth) common factor for almost all of the items. As shown in Fig. 1 (goodness of fit indices for the models and path-diagram for the final five-factor solution), factorial loadings were above 0.30 (between 0.37 and 0.83), meaning that all of the items had a relevant effect on the construct validity of the respective factors. Factors were subsequently labeled according to the original theoretical model.

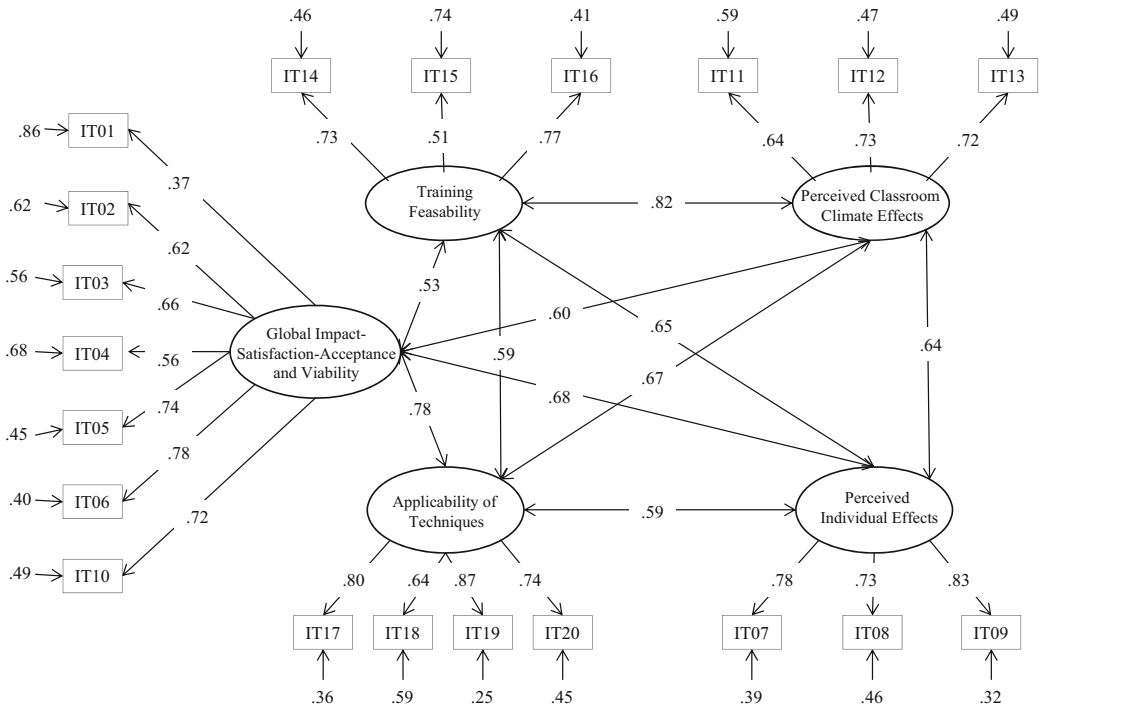
Descriptive statistics for items and factors are showed in Table 1. Mean scores ranged between 3.08 (item 11) and 4.27 (item 13), and all items and factors were symmetrically distributed as the skewness was between $\pm$ 1.5. Finally, in all cases, the correlation item-factor was above 0.30, which implies a relevant contribution of the item to the reliability of its factor.

Regarding reliability and AVE, results showed good internal consistency across all factors: F1 ($\alpha = 0.82$; CR = 0.83; AVE = 0.42), F2 ($\alpha = 0.82$; CR = 0.82; AVE = 0.61), F3 ($\alpha = 0.73$; CR = 0.74; AVE = 0.49), F4 ($\alpha = 0.70$; CR = 0.72; AVE = 0.46), and F5 ($\alpha = 0.84$; CR = 0.85; AVE = 0.59). The overall reliability of the scale was good ($\alpha = 0.84$).

Results relating to discriminant validity are shown in Table 2. *Perceived Classroom Climate Effectiveness* (F3) and *Training Feasibility* (F4) did not demonstrate discriminant validity as their correlation ($r = .82$) was higher than the root square of the AVE (*F3*: 0.70; *F4*: 0.68). *Global Impact-Satisfaction, Acceptance and Viability* (F1) showed a partial overlap with *Perceived Individual Effectiveness* (F2) and *Applicability of Techniques* (F5), as the AVE of the F1 was lower than the correlation with the variables F2 and F5. However, the AVE of both F2 and F5 was higher than their respective correlations with F1.

Results regarding convergent validity are detailed in Table 3. In general, the five factors of the EVSPM-A correlated positively and significantly with the CBCC, CBHRME, EREMIND, EHERM-A, and closed-eyes test. *Global Impact-Satisfaction, Acceptance and Viability* (F1) was the factor with the highest significant correlation coefficients. Furthermore, it was the only factor that was associated with disciplinary sanctions and academic level.

The predictive validity of the scale on REMIND, as measured using the EHERM-A after controlling for sex and academic level, was considerable ($r = .49$; $p < .001$) and related to three of the five factors: F1 (*Global Impact-Satisfaction, Acceptance and Viability*) [$\beta = 0.29$; t (325) = 4.45; $p < .001$],



Goodness-of-Fit Indices of both one-factor and five-factors models

One Factor: χ^2/df= 6.26; GFI = .95; CFI = .92; NNFI = .91; RMSEA = .111 (90% CI: .105 - .118); SRMR = .082

Five Factors: χ^2/df= 2.61; GFI = .98; CFI = .98; NNFI = .97; RMSEA = .062 (90% CI: .055 - .069); SRMR = .058

Bifactorial Model: χ^2/df= 0.59; GFI = .99; CFI = 1.00; NNFI = 1.01; RMSEA = .000 (90% CI: .000 - .000); SRMR = .102

Fig. 1 Path diagram: fit indices and confirmatory factor analysis for the five-factor solution

Table 1. Descriptive statistics (mean, standard deviation, asymmetry, and correlation between item and total scale) of items and factors of the *EVSPM-A*.

Item/factor	*M*	*SD*	*Sk*	Correlat. item-total
Global Impact-Satisfaction, Acceptance, and Viability	4.10	0.59	− 1.21	–
1. I liked to do relaxation-mindfulness before starting some classes.	4.34	0.77	− 1.14	.58
3. Overall, I have managed to relax.	3.94	0.88	− 0.98	.67
7. I know more about mindfulness than before.	3.81	0.92	− 0.79	.47
9. I found it easy in general.	4.04	0.95	− 1.04	.35
13. I think it's useful to do relaxation-mindfulness before class.	4.27	0.99	− 1.43	.55
15. I have been interested in doing relaxation-mindfulness and participated with positive attitude.	4.21	0.82	− 1.24	.43
18. I usually felt calmer after doing relaxation-mindfulness	4.10	1.00	− 1.26	.61
Perceived Individual Effectiveness	3.76	0.80	− 0.88	–
2. I had more energy in class after practicing relaxation-mindfulness	3.53	1.04	− 0.41	.69
10. After making relaxation-mindfulness, I was more excited and happy.	3.69	0.96	− 0.55	.60
12. I had more concentration after doing relaxation-mindfulness.	3.74	0.93	− 0.76	.65
Perceived Classroom Climate	3.71	0.67	− 0.43	–
4. There is more silence in the classroom after we practice relaxation-mindfulness.	3.77	0.89	− 0.54	.54
14. After doing relaxation-mindfulness the atmosphere in class is more enjoyable.	3.89	0.75	− 0.51	.56
17. Improved relationship with teacher who taught it.	3.47	0.92	− 0.30	.38
Training Feasibility	3.60	0.75	− 0.39	–
5. The time devoted to each practice was adequate.	3.83	0.92	− 0.62	.53
11. The number of times (sessions) we have done is enough.	3.08	1.04	− 0.14	.40
19. I am satisfied with the training style of the teacher.	3.89	0.93	− 0.83	.51
Applicability of Techniques	3.83	0.60	− 0.67	–
6. I know enough about each relaxation-mindfulness techniques.	3.96	0.71	− 0.90	.71
8. Techniques are easy to be learned.	3.76	0.71	− 0.55	.66
16. We practiced the number of sessions planned.	3.76	0.78	− 0.74	.66
20. I am satisfied with the proposed techniques.	3.82	0.70	− 0.64	.74

M mean; *SD* standard deviation; *Sk.* skewness. The score of each factor corresponds to the average of the items that compose it

F2 (*Perceived Individual Effectiveness*) [$\beta = 0.16$; t (325) = 2.61, $p = .009$], and F5 (*Applicability of Techniques*) [$\beta = 1.49$; t (325) = 2.62, $p = .009$]. F3 (*Perceived Classroom Climate*) and F4 (*Training Feasibility*) showed no predictive value.

As for incremental validity, in the first regression model, a statistically significant change was observed ($\Delta R^2 = .17$, $p < .001$; $F(2,297) = 20.01$; $p < .001$). When verifying the zero, partial, and semi-partial correlations, a greater magnitude was observed in this predictor, which showed less fluctuation and a considerable semi-partial correlation [$\beta = .44$; $t(297) = 8.157$; $p < .001$; $r_{sp} = .41$]. According to Husley and Meyer (2003), a semi-partial correlation of 0.15 to 0.20 can be considered as good. In the second regression model, a statistically significant change was observed ($\Delta R^2 = .20$; $F(2,291) = 12.11$; $p < .001$). When analyzing which of the factors contributed to the prediction increase, it was observed that factor 1—*Global Impact-Satisfaction, Acceptance and Viability* ($\beta = .30$; $t(291) = 4.07$; $p < .001$; $r_{sp} = .21$) and factor 5—*Applicability of Techniques*

Table 2 Discriminant validity of the *EVSPM-A* factors

	F1	F2	F3	F4	F5		F1–F2	F1–F3	F1–F4	F1–F5	F2–F3	F2–F4	F2–F5	F3–F4	F3–F5	F4–F5
R.S.-AVE	0.65	0.78	0.70	0.68	0.77	Correlation coefficient	0.78	0.60	0.53	0.68	0.67	0.59	0.59	0.82	0.64	0.65

R.S-AVE root square of the AVE, *F1* global impact-satisfaction, acceptance and viability, *F2* perceived individual effectiveness, *F3* perceived classroom climate effectiveness, *F4* training feasibility, *F5* applicability of techniques

Table 3 Correlation coefficients (Pearson's *r*) among the variables

	Global impact(F1)	Perceived individual effectiveness (F2)	Perceived classroom climate (F3)	Training feasibility(F4)	Applicability techniques(F5)	CBCC	EVSPM-A	EREMIND-A	EHERMA	Closed-eyes test	Penalties
F2	.63***	–									
F3	.27***	.37***	–								
F4	.26***	.32***	.53***	–							
F5	.48***	.37***	.30***	.30***	–						
CBCC	.32***	.30***	.17**	.11*	.22***	–					
EVSPM-A	.81***	.80***	.63***	.62	.68***	.33***	–				
EREMIND-A	.39***	.34***	.18**	.11*	.16**	.39***	.35***	–			
EHERMA	.47***	.42***	.23***	.19***	.36***	.18**	.49***	.28***	–		
Closed-Eyes Test	.25***	.09	.02	.03	.15*	.14*	.16**	.18**	.28***	–	
Penalties	–.22***	–.08	–.06	–.03	–.02	–.12*	–.13*	–.15**	–.13*	- .23***	–
CBHRME	.15**	.21***	.02	.04	.05	.24***	.14*	.47***	.14*	.01	.02

*p < .05; *p < .01; ***p < .001

($\beta = .15$; $t(291) = 2.60$; $p = .010$; $r_{sp} = .13$) showed significant values and a lower variability among the correlation types (Table 4).

Analyses of differences by gender and academic level did not show a significant multivariate interaction effect, $F(20, 1304) = 0.74$, $p = .783$. When analyzing gender and academic level (course group) separately, gender did not show a significant multivariate effect, $F(5, 323) = 1.46$, $p = .204$. However, in the case of academic level, a multivariate significant effect was attained, $F(20, 1304) = 4.05$, $p < .001$, $\eta^2 = .06$. As shown in Table 5, the univariate analyses demonstrated that there were significant differences in *Classroom Climate*, *Training Feasibility*, and *Applicability of Techniques*. More specifically, the youngest participants scored lower than the other groups.

Discussion

This study demonstrates that the psychometric properties of the EVSPM-A are good and that the scale presents a clear factorial structure as well as good internal consistency and validity. The obtained validation coefficients likewise strengthen the overall suitability of the scale for assessing SV in MBEPs and MBIs delivered to adolescents in Spanish educational settings.

The observed factorial structure was shown to align with the main authors' theoretical model of SV implicit within the TREVA Program (López-González 2010; López-González et al. 2016a). Of particular note is factor 1 (Global Impact-Satisfaction, Acceptance and Viability) that demonstrated the highest correlation coefficients overall and was the only factor that correlated with discipline and academic performance. It is plausible that students with better behavior and academic acumen are more open to new techniques, such as mindfulness, which can influence their academic performance. Similarly, adolescent students who are less focused on their studies may reject approaches that require additional effort from their part. However, it should be noted that the observations relating to factor 1 could, in part, have arisen due to the broad aspect of this factor that comprises a range of sub-dimensions.

The predictive validity of the EVSPM-A in respect of various assessment variables (e.g., relaxation and mindfulness) supports the idea that SV taps into a broad range of psychotherapeutic change and acceptability outcomes. Findings from this study also indicate that the EVSPM-A can enrich other, more efficacy-focused, outcome data (Barret et al. 2001; Marchant et al. 2013). Indeed, the validation of this scale is in line with the view of researchers who advocate the need for evaluating SV as a complement to other quantitative assessments (Luiselli et al. 2017; Worthen and Luiselli 2017). This study also supports research that advocates evaluating the wider impact and accessibility of MBIs, including the viability

Table 4 Hierarchical regression predicting *EHERM-A* for the analysis of the incremental validity of the *EVSPM-A*

Model 1	*B*	*SE B*	β	*t*	r_{sp}
Step 1					
Gender	−0.07	.08	−.53	−.90	−.05
Grade	0.02	.03	.58	.65	.04
Step 2					
Gender	−.01	.07	−.01	−.13	−.01
Grade	.05	.03	.10	1.70	.09
EREMIND-A	.46	.11	.26	4.03***	.22
CBHRME	.11	.12	.07	.09	.05
Step 3					
Gender	−.05	.07	−.04	−.80	−.04
Grade	.03	.03	.70	1.21	.06
EREMIND-A	.16	.1	.09	1.55	.07
CBHRME	.10	.1	.06	.88	.04
EVSPM-A	.60	.07	.44	8.15***	.41
Model 2					
Step 3 (steps 1 and 2 such as model 1)					
Gender	−.07	.07	−.05	−1.07	−.05
Grade	.04	.03	.08	1.33	.07
EREMIND-A	.12	.11	.07	1.06	.05
CBHRME	.09	.12	.05	.72	.03
Global impact satisfaction	.33	.08	.30	4.07***	.20
Individual effect	.09	.06	.10	1.50	.07
Classroom climate effective	.38	.06	.04	.60	.03
Training feasibility	.01	.05	.01	.06	.00
Applicability of techniques	.17	.06	.15	2.60**	.13

Social Validity Scale of Mindfulness-Based Interventions for Adolescents: Model: 1 R^2 = .004 (ns) for step 1; ΔR^2 = .08 (p = .000) for step 2; ΔR^2 = .17 (p = .000) for step 3. Model 2: R^2 = .005 (ns) for step 1; ΔR^2 = .08 (p = .000) for step 2; ΔR^2 = .20 (p = .000) for step 3]

B non-standardized beta, *SE B* standard error for beta, β standardized beta, *t* *t* test for the beta effect, r_{sp} semi-partial correlation coefficient

*p < .05; *p < .01; ***p < .001.

Table 5 Univariate differences by academic level in each factor of the *EVSPM-A*

	Course group										*F*	η^2
	1st ESO (n = 38)		2st ESO (n = 43)		3th ESO (n = 35)		4th ESO (n = 37)		1st Baccalaureate (n = 16)			
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*		
F1	4.19	0.61	3.99	0.69	4.17	0.64	4.15	0.45	4.03	0.42	1.64	.02
F2	3.92	0.74	3.69	1.00	3.72	0.86	3.81	0.65	3.65	0.61	1.22	.02
F3	3.54c,d	0.59	3.52c,d	0.76	3.96a,b	0.64	3.88a,b	0.65	3.77	0.44	6.98**	.08
F4	3.50c	0.65	3.63	0.82	3.86a,d	0.67	3.49c	0.89	3.50	0.49	3.29*	.04
F5	3.79	0.66	3.75	0.63	3.72d	0.67	4.00c	0.49	3.92	0.42	2.63*	.03

Differences by academic level, being 1° ESO (a), 2° ESO (b), 3° ESO (c), 4° ESO (d) and 1° Baccalaureate (e), according to the Bonferroni Post Hoc test

F1 global impact-satisfaction, acceptance and viability, *F2* individual effect, *F3* classroom climate effects, *F4* training feasibility, *F5* applicability of techniques, *ESO* Compulsory Secondary Education

*p < .05; **p < .001

of instruction and teaching techniques (Sainio et al. 2018; Crane et al. 2013).

More specifically, the EVSPM-A appears to be an effective and tailored means of evaluating SV in MBEPs, where adolescents undergoing mindfulness training can experience perceptual changes in terms of how participants relate to and evaluate a given interventional approach (Sapthiang et al. 2019). It is for this reason that the EVSPM-A, an MBEP-specific measure of SV, should lead to a more accurate appraisal of the wider utility of MBEPs among adolescents and other key intervention stakeholders (Metz et al. 2013). The 20-item EVSPM-A may also open possibilities to reduce the number of self-report questionnaires used in MBEP evaluation studies, which can be burdensome to participants when used excessively (Bergomi et al. 2013; Kuby et al. 2015).

The measurement of SV using the EVSPM-A should help in the identification of integration issues early on in the development and evaluation of the MBEPs, including key intervention assessment considerations such as cost-effectiveness, intervention integrity, competence of the trainer, training model suitability, and adherence to practice (Crane and Hecht 2018). Furthermore, since the scale is inspired by the TREVA program which draws on design factors used in both MBSR-based approaches as well as second-generation mindfulness-based interventions (Van Gordon et al. 2015), the EVSPM should have utility for evaluating SV across a broad range of MBEPs and could contribute in the transition of MBIs from clinical settings to school settings.

Limitations and Future Research

Although the present study yielded additional support for the validity of the EVSPM-A, findings should be viewed in light of the study limitations. Firstly, discriminant validity was not observed for all factors of the scale when viewed separately. However, it should be noted that discriminant validity was partial among some pairs of factors, but logical and sustainable from a theoretical point of view. Secondly, the fact that there is a degree of overlap in meanings between some of the factors could limit internal consistency. Thirdly, the EFA and CFA were carried out with two partially overlapping samples (CFA was conducted with the same sample as EFA after including 92 extra participants). Therefore, in future research, it would be useful to investigate if the current results replicate across different samples. Finally, the sample only included students from three private and public Spanish education centers, meaning that the generalizability of findings to specific student groups (e.g., special needs students) or non-Spanish students needs to be explored further.

In terms of future directions, additional studies are recommended to further evaluate the psychometric properties and utility of the EVSPM-A. It would also be useful to validate the scale in other languages as well as develop adaptations for use with different adolescent age groups (i.e., 12–14 years—early adolescence, and 15 to 18). Finally, it would be appropriate to carry out some comparative studies of the EVSPM-A with different qualitative measures to reinforce its validity and reliability.

The present research represents a promising development in the evaluation of SV in MBEPs. The EVSPM-A has acceptable reliability and validity for evaluating the utility and wider acceptability of mindfulness interventions and programs in educational settings. A lack of research into the wider effects—that are often implied (i.e., but not empirically supported) in MBEP and MBI intervention studies (Eklund et al. 2017)—makes the EVSPM-A particularly valuable. More specifically, the EVSPM-A can complement other types of measurement that are mainly focused on assessing a small number of pre-defined intervention effects.

Acknowledgments We wish to express our deepest appreciation to the individuals who voluntarily participated in the project and university employees for their support with the study.

Author's Contributions LLG and AA conceived and designed the study and wrote and edited all the different versions of the ms and the final paper. DHF conducted data analyses. TS participated in recruiting the sample, and in the administration of the different instruments. WVG contributed to the writing and final editing of the manuscript.

Compliance with Ethical Standards

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The study was approved by the Bioethics Committee of the University of Barcelona (Spain).

Conflict of Interest The authors declare that they have no conflict of interest.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

Ager, K., Albrecht, N. J., & Cohen, M. (2015). Mindfulness in schools research project: Exploring Student's perspectives of mindfulness—what are students' perspectives of learning mindfulness practices at school. *Psychology, 6*(7), 896–914. https://doi.org/10.4236/psych.2015.67088.

Baltzell, A. (2016). *Mindfulness and performance*. London: Cambridge University Press.

Barret, P. M., Short, A. L., & Wescombe, K. (2001). Examining the social validity of the FRIENDS treatment program for anxious children. *Behaviour Change, 18*(2), 63–67.

Beauchemin, J., Hutchins, T. L., & Patterson, F. (2008). Mindfulness meditation may lessen anxiety, promote social skills, and improve

academic performance among adolescents with learning disabilities. *Complementary Health Practice Review, 13*(1), 34–45. https://doi.org/10.1177/1533210107311624.

Bergomi, C., Tschacher, W., & Kupper, Z. (2013). The assessment of mindfulness with self-report measures: Existing scales and open issues. *Mindfulness, 4*(3), 191–202. https://doi.org/10.1007/s12671-012-0110-9.

Bluth, K., & Blanton, P. W. (2015). The influence of self-compassion on emotional well-being among early and older adolescent males and females. *The Journal of Positive Psychology, 10*(3), 219–230. https://doi.org/10.1080/17439760.2014.936967.

Bollen, K. A., & Curran, P. J. (2005). *Latent curve models: a structural equation perspective*. Hoboken, NJ: Wiley.

Carmines, E. G., & McIver, J. P. (1981). Analyzing models with unobserved variables: analysis of covariance structures. In G. W. Bornstedt & E. F. Borgatta (Eds.), *Social Measurement: Current issues* (pp. 112–133). Beverly Hills: Sage.

Carroll, R., & St. Peter, C. C. (2014). Methods for assessing social validity of behavioral intervention plans for children with attention deficit hyperactivity disorder. *Acta de Investigación Psicológica, 4*(3), 1642–1656. https://doi.org/10.1016/S2007-4719(14)70971-8.

Carter, S. L. (2009). *The social validity manual: a guide to subjective evaluation of behavior interventions*. London: Academic Press.

Crane, R. S., & Hecht, F. M. (2018). Intervention integrity in mindfulness-based research. *Mindfulness, 9*, 1370–1380. https://doi.org/10.1007/s12671-018-0886-3.

Crane, R. S., Eames, C., Kuyken, W., Hastings, R. P., Williams, J. M. G., Bartley, T., et al. (2013). Development and validation of the mindfulness-based interventions–teaching assessment criteria (MBI: TAC). *Assessment, 20*(6), 681–688. https://doi.org/10.1177/1073191113490790.

Da Fonceca, V. (1998). *Manual de Observación Psicomotriz*. Zaragoza: Inde.

Dariotis, J., Miralbal-Beltran, R., Cluxtin-Keller, F., Gould, L. F., Greenberg, M. T., & Mendelson, M. (2016). A qualitative evaluation of student learning and skills use in a school-based mindfulness and yoga program. *Mindfulness, 7*(1), 76-89. https://doi.org/10.1007/s12671-015-0463-y.

Davidson, R. J., Dunne, J., Eccles, J. S., Engle, A., Greenberg, M., Jennings, P., & Vago, D. (2012). Contemplative practices and mental training: prospects for American education. *Child Development Perspectives, 6*(2), 146–153. https://doi.org/10.1111/j.1750-8606.2012.00240.x.

Eklund, K., O'malley, M., & Meyer, L. (2017). Gauging mindfulness in children and youth: school-based applications. *Psychology in the Schools, 54*(1), 101–114. https://doi.org/10.1002/pits.21983.

Emerson, L. M., Leyland, A., Hudson, K., Rowse, G., Hanley, P., & Hugh-Jones, S. (2017). Teaching mindfulness to teachers: a systematic review and narrative synthesis. *Mindfulness*, *8(5)*, 1136-1149. https://doi.org/10.1007/s12671-017-0691-4.

Felver, J. C., Celis-de Hoyos, C. E., Tezanos, K., & Singh, N. N. (2016). A systematic review of mindfulness-based interventions for youth in school settings. *Mindfulness, 7*(1), 34–45. https://doi.org/10.1007/s12671-015-0389-4.

Flook, L., Smalley, S. L., Kitil, M. J., Galla, B. M., Kaiser-Greenland, S., & Locke, J. (2010). Effects of mindful awareness practices on executive functions in elementary school children. *Journal of Applied School Psychology, 26*, 70–95. https://doi.org/10.1080/15377900903379125.

Forman, K., & Soloff, G. (2014). *Preparing for educational leadership* (2nd ed.). New York: Pearson Education.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 27*, 39–50.

Greenberg, M. T., & Harris, A. R. (2012). Nurturing mindfulness in children and youth: Current state of research. *Child Development Perspectives, 6*, 161–166. https://doi.org/10.1111/j.1750-8606.2011.00215.x.

Hanley, G. P. (2012). Functional assessment of problem behavior: dispelling myths, overcoming implementation obstacles, and developing new lore. *Behavior Analysis in Practice, 5*(1), 54–72. https://doi.org/10.1007/BF03391818.

Himelstein, S. (2016). *A mindfulness-based approach to working with high-risk adolescents*. London: Routledge.

Hokke, H., Lievense, P., & Sterkenburg, P. S. (2014). The effects of mindfulness for people with visual and moderate intellectual disability and their caregivers. *Journal of Applied Research in Intellectual Disabilities, 27*(4), 311.

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

Husley, J., & Meyer, G. (2003). The incremental validity of psychological testing and assessment: conceptual, methodological, and statistical issues. *Psychological Assessment, 15*, 466–455. https://doi.org/10.1037/1040-3590.15.4.446.

Jennings, P. A., Brown, J. L., Frank, J. L., Doyle, S., Oh, Y., Davis, R., Rasheed, D., De Weese, A., DeMauro, A. A., Cham, H., & Greenberg, M. T. (2017). Impacts of the CARE for Teachers Program on teachers' social and emotional competence and classroom interactions. *Journal of Educational Psychology. Advance online publication.* https://doi.org/10.1037/edu0000187.

Jöreskog, K., & Sörbom, D. (1997). *Lisrel 8: user's reference guide*. Lincolnwood: Scientific Software International.

Kallapiran, K., Koo, S., Kirubakaran, R., & Hancock, K. (2015). Review: effectiveness of mindfulness in improving mental health symptoms of children and adolescents: a meta-analysis. *Child and Adolescent Mental Health, 20*(4), 182–194. https://doi.org/10.1111/camh.12113.

Kern, L., & Manz, P. (2004). A look at current validity issues of school-wide behavior support. *Behavioral Disorders, 30*, 47–59.

Klingbeil, D. A., Renshaw, T. L., Willenbrink, J. B., Copek, R. A., Chan, K. T., Haddock, A., et al. (2017a). Mindfulness-based interventions with youth: a comprehensive meta-analysis of group-design studies. *Journal of School Psychology, 63*, 77–103. https://doi.org/10.1016/j.jsp.2017.03.006.

Klingbeil, D. A., Fischer, A. J., Renshaw, T. L., Bloomfield, B. S., Polakoff, B., Willenbrink, J. B., et al. (2017b). Effects of mindfulness-based interventions on disruptive behavior: a meta-analysis of single-case research. *Psychology in the Schools, 54*(1), 70–87. https://doi.org/10.1002/pits.21982.

Kuby, A. K., McLean, N., & Allen, K. (2015). Validation of the child and adolescent mindfulness measure (CAMM) with non-clinical adolescents. *Mindfulness, 6*(6), 1448–1455. https://doi.org/10.1007/s12671-015-0418-3.

Kuyken, W., Weare, K., Ukoumunne, O.C Vicary, R., Motton, N., Burnett, R., Cullen, C., Hennelly, S., & Huppert, F. (2013). Effectiveness of the mindfulness in schools programme: non-randomised controlled feasibility study. The British Journal of Psychiatry, 203(2), 126–131; doi: https://doi.org/10.1192/bjp.bp.113.126649.

Leonard, N. R., Jha, A. P., Casarjian, B., Goolsarran, B., Garcia, C., Cleland, C. M., Gwadz, M. V., & Massey, Z. (2013). Mindfulness training improves attentional task performance in incarcerated youth: a group randomized controlled intervention trial. *Frontiers in Psychology, 4*(792), 1–10. https://doi.org/10.3389/fpsyg.2013.00792.

Liehr, P., & Diaz, N. (2010). A pilot study examining the effect of mindfulness on depression and anxiety for minority children. *Archives of Psychiatric Nursing, 24*(1), 69–71. https://doi.org/10.1016/j.apnu.2009.10.001.

López-González, L. (2010). *Disseny i desenvolupament d'un programa de relaxació vivencial aplicada a l'aula*. Doctoral thesis, University of Barcelona - (UB). Barcelona, Spain. Available at: http://tdx.cat/bitstream/handle/10803/2364/LLG_TESI.pdf?sequence=1. Accessed 19 March 2019.

López-González, L., & Bisquerra, R. (2013). Validación y análisis de una escala breve para evaluar el clima de clase en Educación Secundaria. *ISEP SCIENCE, 5*, 62–77.

López-González, L., Álvarez-González, M., & Bisquerra, R. (2016a). Mindfulness and action research in secondary schools. Gestation of the TREVA program. *Revista Interuniversitaria de Formación del Profesorado, 87*(30,3), 75–91.

López-González, L., Amutio, A., Herrero-Fernández, D., & Bisquerra, R. (2016b). Validation of a relaxation-mindfulness skills and states scale for adolescents. *Revista Interuniversitaria de Formación del Profesorado, 87*(30,3), 93–105.

López-González, L., Amutio, A., Oriol, X., & Bisquerra, R. (2016c). Habits related to relaxation and mindfulness of high school students: influence on classroom climate and academic performance. *Journal of Psychodidactics, 21*(1), 121.138. https://doi.org/10.1387/RevPsicodidact.13866.

López-González, L., Amutio, A., & Herrero-Fernández, D. (2018a). The relaxation-mindfulness competence of secondary and high school students and its influence on classroom climate and academic performance. *European Journal of Education and Psychology, 11*(1), 5–17.

López-González, L., Amutio, A., Oriol, X., Gázquez, J. J., Pérez-Fuentes, M. C., & Molero, M. M. (2018b). Development and validation of the relaxation-mindfulness scale for adolescents (EREMIND-A). *Psicothema, 30*(2), 224–231.

Lorenzo-Seva, U., & Ferrando, P. J. (2013). Factor 9.2: a comprehensive program for fitting exploratory and Semiconfirmatory factor analysis and IRT models. *Applied Psychological Measurement, 37*(6), 497–498.

Luiselli, J. K., Worthen, D., Carbonell, L., & Queen, A. H. (2017). Social validity assessment of mindfulness education and practices among high school students. *Journal of Applied School Psychology, 33*(2), 124–135. https://doi.org/10.1080/15377903.2016.1264531.

Marchant, M., Heath, M., & Miramontes, N. Y. (2013). Merging empiricism and humanism role of social validity in the school-wide positive behavior support model. *Journal of Positive Behavior Interventions, 15*(4), 221–230. https://doi.org/10.1177/1098300712459356.

Mendelson, T., Greenberg, M. T., Dariotis, J. K., Gould, L. F., Rhoades, B. L., & Leaf, P. J. (2010). Feasibility and preliminary outcomes of a school-based mindfulness intervention for urban youth. *Journal of Abnormal Child Psychology, 38*, 985–994. https://doi.org/10.1007/s10802-010-9418-x.

Metz, S. M., Frank, J. L., Reibel, D., Cantrell, T., Sanders, R., & Broderick, P. C. (2013). The effectiveness of the learning to BREATHE program on adolescent emotion regulation. *Research in Human Development, 10*(3), 252–272. https://doi.org/10.1080/15427609.2013.818488.

Mîndrilă, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: a comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society, 1*(1), 60–66.

Pallozzi, R., Wertheim, E., Paxton, S., & Ong, B. (2016). Trait mindfulness measures for use with adolescents: a systematic review. *Mindfulness*. https://doi.org/10.1007/s12671-016-0567-z.

Parker, A. E., & Kupersmidt, J. B. (2016). Two universal mindfulness education programs for elementary and middle-school students: master mind and moment. In K. Schonert-Reichl & R. Roeser (Eds.), *Handbook of mindfulness in education. Mindfulness in Behavioral Health*. New York, NY: Springer. https://doi.org/10.1007/978-1-4939-3506-2_21.

Rennekamp, R. A., & Arnold, M. E. (2009). What progress, program evaluation? Reflections on a quarter-century of extension evaluation practice. *Journal of Extension [On-line], 47*(3), Article 3COM1 Available at: https://www.joe.org/joe/2009june/pdf/JOE_v47_3comm1.pdf. Accessed 19 March 2019.

Sainio, M., Herkama, S., Turunen, T., Rönkkö, M., Kontio, M., Poskiparta, E., & Salmivalli, C. (2018). Sustainable antibullying program implementation: school profiles and predictors. *Scandinavian Journal of Psychology*. https://doi.org/10.1111/sjop.12487.

Saltzman, A., & Goldin, P. (2008). Mindfulness-based stress reduction for school-age children. In L. A. Greco & S. C. Hayes (Eds.), *Acceptance and mindfulness treatments for children and adolescents: a practitioner's guide* (pp. 139–161). Oakland, CA, US: New Harbinger Publications.

Sanger, K. L., & Dorjee, D. (2015). Mindfulness training for adolescents: a neurodevelopmental perspective on investigating modifications in attention and emotion regulation using event-related brain potentials. *Cognitive, Affective, & Behavioral Neuroscience, 15*(3), 696–711. https://doi.org/10.3758/s13415-015-0354-7.

Sapthiang, S., Van Gordon, W., & Shonin, E. (2019). Mindfulness in schools: a health promotion approach to improving adolescent mental health. *International Journal of Mental Health and Addiction, 17*, 112-119. https://doi.org/10.1007/s11469-018-0001-y.

Schonert-Reichl, K. A., & Roeser, R. W. (2016). *Handbook of mindfulness in education*. New York: Springer.

Semple, R. J., Droutman, V., & Reid, B. A. (2017). Mindfulness goes to school: things learned (so far) from research and real-world experiences. *Psychology in the Schools, 54*(1), 29–52. https://doi.org/10.1002/pits.21981.

Shababi, S. (2014). *Use of a mindfulness practice to decrease problem behavior and increase engaged time of three students in an elementary school setting*. Vancouver: University of British Columbia.

Sheinman, N., Levy, S., & Lahav, A. (2011). *Mindfulness for a whole school–mindfulness and mind-body skills with children*. Unpublished manuscript.

Shonani, A., & Slone, M. (2013). Middle school transition from the strengths perspective: young adolescents' character strengths, subjective well-being, and school adjustment. *Journal of Happiness Studies, 14*, 1163–1181. https://doi.org/10.1007/s10902-012-9374-y.

Shonin, E., Van Gordon, W., & Griffiths, M. D. (2012). The health benefits of mindfulness-based interventions for children and adolescents. *Education and Health, 30*, 94–97.

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences, 42*(5), 893–898. https://doi.org/10.1016/j.paid.2006.09.017.

Stone, J. R. A. (2014). *Mindfulness in schools: taking present practice into account*. DECP Debate Edition. Available at: https://northampton.academia.edu/JacquelineStone/Papers. Accessed 19 March 2019.

Tan, L. B. G. (2016). A critical review of adolescent mindfulness-based programmes. *Clinical Child Psychology and Psychiatry, 21*(2), 175–177. https://doi.org/10.1177/1359104515577486.

Van Gordon, W., Shonin, E., & Griffiths, M. D. (2015). Towards a second-generation of mindfulness-based interventions. *Australia and New Zealand Journal of Psychiatry, 49*, 591–591.

Wiesner, M., & Schanding, G. T. (2013). Exploratory structural equation modeling, bifactor models, and standard confirmatory factor analysis models: application to the BASC-2 behavioral and emotional screening system teacher form. *Journal of School Psychology, 51*, 751–763. https://doi.org/10.1016/j.jsp.2013.09.001.

Wisner, B. (2017). *Mindfulness and meditation for adolescents*. London: Springer.

Wolf, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of*

Applied Behavior Analysis, 11, 203–214. https://doi.org/10.1901/2Fjaba.1978.11-203.

Worthen, D., & Luiselli, J. K. (2017). Social validity assessment and intervention evaluation of mindfulness education and practices with high school students. *Mindfulness, 8*(4), 903–910. https://doi.org/10.1007/s12671-016-0664-z.

Zenner, C., Herrnleben-Kurz, S., & Walach, H. (2014). Mindfulness-based interventions in schools—a systematic review and meta-analysis. *Frontiers in Psychology, 5*, 603. https://doi.org/10.3389/fpsyg.2014.00603.