



Meta-analysis of Reliability and Validity of the Bergen Social Media Addiction Scale (BSMAS)

Rossella Bottaro¹ · Mark D. Griffiths² · Palmira Faraci¹

Accepted: 22 February 2025

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

Abstract

The present meta-analysis reviewed and summarized the psychometric properties of the Bergen Social Media Addiction Scale (BSMAS), the most widely used tool for assessing social media addiction (SMA) in research and clinical practice. Following the PRISMA 2020 Statement guidelines, seven databases (*PubMed*, *PsycArticles*, *PsycInfo*, *Psychology and Behavioral Sciences Collection*, *Medline*, *Wiley Online Library*, and *Web of Science*) were searched for studies reporting the dimensionality, item characteristics, reliability, and validity of the BSMAS. A total of 28 studies ($N=62,406$) were reviewed. The unidimensionality of the BSMAS was unanimously confirmed with an optimal pooled Cronbach's alpha coefficient (0.83). Likewise, the pooled association between the BSMAS and anxiety, depression, internet gaming disorder, and stress supported its construct validity. There was preliminary and encouraging evidence for other related measures and criteria, and test-retest reliability, although these were qualitatively evaluated due to the limited number of studies. Pending common nosographic categorization, the meta-analytic findings support the appropriateness and validity of conclusions regarding SMA reached using the BSMAS. Further evidence-based, randomized studies targeting various populations and subgroups are warranted.

Keywords Bergen Social Media Addiction Scale · Meta-analysis · Psychometrics · Reliability · Validity

Introduction

Although not yet included in current official diagnostic manuals, social media addiction (SMA) is increasingly recognized as a mental health disorder. The growing interest by scholars is related to the proliferation of social networking sites (SNSs). There were 5.04 billion social profiles active in 2024, representing more than 62% of the world's population, who spend 2.23 hours a day using social media, mainly to stay in touch with friends and family (We Are Social & Meltwater, 2024).

SMA has been defined as being excessively concerned, overly preoccupied with social media, and spending considerable time and energy on social media that it interferes with an individual's ability to engage in social activities, form interpersonal relationships, study or work, or maintain health and wellbeing (Andreassen & Pallesen, 2014). However, a recent review of theories and models regarding SMA (Sun & Zhang, 2021) highlighted those terms such as 'problematic social media use', 'compulsive social media use' or 'addictive social media use' are often used interchangeably to refer to this type of behavior. Moreover, 25 different theories/models on the topic were identified, highlighting multiple comorbidities with psychiatric disorders such as anxiety and depression (Sun & Zhang, 2021; Szczygieł & Podwalski, 2020; Wang et al., 2022). Therefore, inconsistencies exist in the current literature regarding the definition and assessment of SMA. As a consequence, researchers and clinicians lack sufficient gold standards to guide their studies and practice.

The Bergen Social Media Addiction Scale (BSMAS; Andreassen et al., 2016) is one of the most commonly used psychometric instruments to assess SMA. The BSMAS is a six-item self-report scale with a five-point Likert scale response form ranging from *very rarely* (1) to *very often* (5) (e.g., "*How often during the last year have you tried to cut down on the use of social media without success?*") with a summed score ranging from 6 to 30. Each item reflects one of the six core criteria in the addiction components model of addiction (i.e., salience, mood modification, tolerance, withdrawal, conflict, and relapse; Griffiths, 2005).

Notably, the BSMAS is an adaptation of the Bergen Facebook Addiction Scale (BFAS, Andreassen et al., 2012) also based on the six aforementioned criteria. In the adaption, the word 'Facebook' was replaced with the words 'social media', where social media is defined in the instructions to participants as "*Facebook, Twitter, Instagram, and the like*". The one-factor structure of the BFAS was also adopted for the BSMAS. However, the scale's unidimensionality was assumed by the developers of the BSMAS although no confirmatory

factor analysis (CFA) was reported in the study that first used it. Although this gap has since been filled by the numerous validations of the BSMAS in different languages, which has led to its widespread use, no meta-analytic evidence of its reliability and validity is currently available in the literature. Therefore, by systematically reviewing and meta-analyzing the published studies to date, the present study's main research questions (RQs) in relation to the psychometric studies that have been carried out internationally were: (i) What are the characteristics of the studies that have psychometrically evaluated the BSMAS? (RQ1); (ii) What is the research quality of the studies that have evaluated the BSMAS? (RQ2); (iii) What is the consistency regarding the dimensionality of the BSMAS? (RQ3); How reliable is the BSMAS? (RQ4), (v) How valid is the BSMAS? and (vi) is there any publication bias in the BSMAS studies.

Methods

A systematic review and meta-analysis of the previously published research on BSMAS was conducted in strict adherence to the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines (Page et al., 2021). It followed the three steps of identification, screening, and coding outlined below. The flow diagram is shown in Figure 1 and the checklist can be found in Supplementary Table S1.

Search strategy

A comprehensive search of seven databases was conducted (i.e., *PubMed*, *PsycARTICLES*, *PsycINFO*, *Psychology and Behavioral Sciences Collection*, *MEDLINE*, *Wiley Online Library*, and *Web of Science*). This search was supplemented through searching other sources (e.g., *Google Scholar*) and by manually searching the reference lists of the included studies. The authors carried out the search in August 2024. The years of the included studies ranged from 2016 (i.e., the year of the first publication using the BSMAS) to 2024. To

identify the studies, the keyword ‘Bergen Social Media Addiction Scale’ was searched for in paper titles and abstracts.

Screening

The following eligibility criteria were used to screen the selected studies: (i) being published in the English language, (ii) quantitative research published in peer-reviewed journals, and (iii) studies that specifically performed psychometric analysis of the BSMAS (e.g., dimensionality, reliability, and/or validity). The exclusion criteria were: (i) previous literature reviews, books, theses, conference papers, and abstracts, (ii) studies that included BSMAS only as a measure of SMA in specific groups or models without examining its psychometric characteristics, (iii) studies that only provided the estimate prevalence of SMA using the BSMAS with the assumption that the reliability and validity of BSMAS had already been established, (iv) studies that only reported Cronbach’s alpha to support the reliability of the scale in their study but did not specifically focus on the psychometric properties of the BSMAS; and (v) studies that psychometrically evaluated the 18-item BSMAS.

The only exception to these exclusion criteria is the original study by the developers of the BSMAS (i.e., Andreassen et al., 2016): While it is not a validation study, given that it was the first study to use the BSMAS and it reported some psychometric properties, it was included due to its relevance to later studies.

Using the inclusion and exclusion criteria, one author first reviewed the titles and abstracts of the papers to determine which might be included. The authors used EndNote (X9; The EndNote Team, 2013) to initially search for the duplicate papers. The reference lists of the included papers were then searched for more potentially eligible studies. The initially selected papers were reviewed by a separate author. There were no differences in the inclusion or exclusion judgments made by the authors.

Coding

The included studies were coded for authors and year of publication, sample size, age of participants, BSMAS language, and online or offline administration procedure. Dimensionality and item properties were summarized by number of factors, method, and factor loadings. All available reliability and validity values were extracted. The authors chose Cronbach's alpha coefficient as the internal consistency index because it was present in all papers; any value ≥ 0.70 was considered acceptable. All Pearson's r coefficients were coded to summarize convergent, discriminant, and criterion validity with related measures.

Data analyses

Firstly, the quality appraisal of the 28 included studies was assessed according to the Cochrane criteria (Higgins et al., 2023). It was plotted using the Risk of Bias VISualization (*robvis*; McGuinness & Higgins, 2021) tool for a generic study (Figure 2-3). Moreover, the present study also used the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) checklist (Mokkink et al., 2010) to evaluate the methodological quality of studies examining the psychometric properties of instruments. According to Schellingerhout et al. (2012), an overall score for the methodological quality of each included study for the main properties (i.e., 10 boxes) is obtained by taking the lowest rating of any item in a box (see <http://www.cosmin.nl>).

After coding the included studies, the information on dimensionality was presented qualitatively due to the absence of discrepancies between studies. Indeed, unidimensionality was unanimously supported despite the use of different methods and factor loadings. Test-retest and item-total correlations were also reported in narrative form when possible. Cronbach's alpha coefficients were chosen to summarize the reliability of the instrument in the present random-effect meta-analysis. For the longitudinal studies (Chen, Strong, et al., 2020; Gomez

et al., 2024; Shin, 2022), the baseline Cronbach's alpha was included. When studies reported two or more Cronbach's alpha values referred to subgroups, the mean value was used due to their similarity (e.g., 0.85 for Hong Kong and 0.82 for Taiwan; Leung et al., 2020).

Pearson's r correlation coefficient between BSMAS and other measures used were considered as reported in the included papers, as well as the effect size. This coefficient was then transformed into a Fisher-z score based on sample size (Lipsey & Wilson, 2001). According to the criteria formulated by Cohen (1988) for effect sizes, $r=0.10$ were considered small, $r=0.30$ medium, and $r=0.50$ large. When the number of coefficients between studies was substantial (i.e., five or more studies), a quantitative meta-analysis was performed (e.g., when the association between SMA and anxiety was studied). However, when the association was examined in fewer than five studies, a qualitative synthesis was reported (Myung, 2023).

Cohen's Q statistics were used to evaluate the heterogeneity within the meta-analysis. The level of heterogeneity was assessed according to the Cochrane guidelines (Higgins et al., 2023), which were as follows: 0-40% nonsignificant, 30-60% moderate, 50-90% substantial, and 75-100% considerable. Based on the current available variables in the literature, moderation analyses were carried out to explain heterogeneity.

Publication bias was tested by performing Begg's and Egger's tests ($p<0.1$, evidence of publication bias) and graphically using a funnel plot (Begg & Mazumdar, 1994; Egger et al., 1997). Moreover, the trim-and-fill method was performed to estimate the impact of potential missing studies due to publication bias in the funnel plot and to adjust the overall effect estimate (Duval & Tweedie, 2000a, 2000b; Shi & Lin, 2019). More specifically, the trim-and-fill method is first *trims* the studies that cause a funnel plot's asymmetry and then *fills* the imputed missing studies in the funnel plot based on the bias-corrected overall estimate to minimize the impact

of publication bias on the overall effect estimate. After assessing heterogeneity, publication bias was evaluated when the summary effect size was available.

Finally, data analysis was performed using the R software and programming language (R Core Team, 2021), using the ‘metafor’ package for a random effects model (Viechtbauer, 2010). The `rma.mv` function was employed for fitting linear (mixed-effects) models to metanalytic multivariate/multilevel fixed-effects and random/mixed-effects models, with or without moderators. Graphically, forest plots showed effect sizes and associated 95% confidence intervals for each study, and the funnel plot served as an estimate of publication bias.

Results

Sample characteristics of studies evaluating the BSMAS (RQ1)

The initial search yielded 165 studies. After eliminating 22 duplicates, 154 papers were screened and 28 were selected based on the established eligibility criteria. The study selection procedure, which was guided by the PRISMA Statement guidelines (Page et al., 2021), is shown in Figure 1.

The psychometric properties of the Chinese BSMAS ($n=5$) and English BSMAS ($n=4$) were the most studied. However, 28 languages were used to validate the BSMAS in different countries, such as Italian, Hungarian, Greek, Spanish, etc. (for more details, see Table 1), using a back-translation procedure. The studies comprised a total of 62,406 participants, with independent sample sizes ranging from 247 to 23,533, and a mean age of 21.89 years ($SD=4.25$). Participants were mostly university students ($n=14$), adolescents ($n=7$), and adults ($n=7$).

Most of the included studies adopted a cross-sectional design except for three longitudinal studies (Table 1), were published in peer reviewed journals, and used the six-item

version of the BSMAS (BSMAS mean score of all included studies: $M=14.99$, $SD=4.52$). Most data were collected through online surveys ($n=17$), others were compiled in-person setting ($n=8$) or both ($n=2$); three studies did not specify this information. In total, 28 effect sizes were identified for Cronbach's alpha (ranging from 0.73 to 0.88). Moreover, internet gaming disorder, anxiety, depression, and stress were the most studied variables in evaluating construct validity.

Research quality assessment of BSMAS studies (RQ2)

According to the Cochrane criteria for the classification of risk of bias, the included studies had a moderate overall risk (Figures 2 and 3). Notably, the highest risk was related to the random sequence generation. The studies mostly used convenience samples and surveys were distributed through social media platforms such as *Facebook*, *Instagram*, *TikTok*, or internal university channels such as emails. The risk of incomplete outcome data raised concerns due to the lack of information about the mandatory response format (i.e., if the authors expected missing data) or as part of a larger research project.

Moreover, the authors assessed the quality of the included psychometric papers according to the COSMIN checklist (for details, see Supplementary Materials S1). As briefly shown in Table 2 and Figure 4, all included studies provided optimal evidence of internal consistency by performing Cronbach's alpha ($\alpha \geq 0.70$). Most of them (26 studies out of 28) tested the construct validity of the BSMAS and evaluated its responsiveness. Moreover, the structural validity was well-tested in 22 of the 28 studies. Other studies found empirical support for the measurement invariance ($n=14$) and criterion validity ($n=11$) of the BSMAS. Finally, the results also highlighted that most of the included studies (22 studies out of 28) provided an adequate description of the BSMAS development process.

Dimensionality and item properties (RQ3)

All included studies supported the unidimensionality of the BSMAS by using the six-items version. The most widely used technique was confirmatory factor analysis (CFA; $n=24$), which showed satisfactory factor loadings on a single latent factor ($\lambda \geq 0.30$). Only one study performed an exploratory factor analysis (EFA; Naher et al., 2022) and two performed a principal component analysis (Brailovskaia, & Margraf, 2022; Žmavc et al., 2022). Satisfactory item-total correlations were found for all items (ranging from 0.49 for Item 4 to 0.73 for Item 2). Using an item response theory (IRT) approach, six studies (El Abiddine et al., 2024; Lin et al., 2017; Naher et al., 2022; Stănculescu, 2023; Tung et al., 2022; Zarate et al., 2023) provided further information about BSMAS items' characteristics, such as adequate discrimination and difficulty parameters for all items (Stănculescu, 2023; Tung et al., 2022), more information at middle and middle-high trait levels (peak ranged from +0.5 to +2.0 SD; Zarate et al., 2023), and no differential item function across gender or time spent on social media (Lin et al., 2017; Naher et al., 2022).

Reliability of the BSMAS(RQ4)

Internal consistency. All of the 28 included studies reported satisfactory Cronbach's alpha values, ranging from 0.73 to 0.88 (Figure 5). Firstly, the Shapiro-Wilk's test suggested that the values were normally distributed ($p > 0.001$). After performing a random-effect meta-analysis, the estimated effect size was 0.83 (standard error=0.01, $p < 0.001$, C.I. 95% [0.81 – 0.85]). However, the findings of the meta-analysis showed a substantial degree of heterogeneity ($I^2=98.53\%$, Cochran's $Q=726.79$, $df=26$, $p < 0.001$).

A moderator analysis of studies by sample size was performed. However, there were no differences between studies ($p=0.07$) and considerable residual heterogeneity ($I^2=97.65\%$, Cochran's $Q=447.62$ $df=26$, $p < 0.001$). Subsequently, population type was also tested (i.e., university students or others) as a moderator which showed nonsignificant results ($p=0.56$) with relevant residual heterogeneity ($I^2=98.42\%$, Cochran's $Q=646.19$, $df=26$, $p < 0.001$).

Finally, the assessment setting (i.e., online vs. offline) was also tested as a moderator. However, there were no differences between the online and offline settings ($p=0.49$), with noticeable residual heterogeneity ($I^2=98.39\%$, Cochran's $Q=482.87$, $df=22$, $p<0.001$).

Test-retest reliability. Test-retest reliability was evaluated in only three studies. As a consequence of their small size and varying time-frames, the results here are presented qualitatively. Shin (2022) and Chen, Strong et al. (2020) showed strong associations at three-week and three-month intervals ($r=0.75$ and 0.72 , respectively). Similarly, Gomez et al. (2024) highlighted moderate reliability at 12 and 24 months ($r=0.53$ and 0.42 , respectively).

Validity of the BSMAS (RQ5)

Construct validity. The validity of the BSMAS validity studied in relation to relevant convergent and divergent constructs (forest plot in Figure 6). For convergent validity, the association with depression ($n=11$), anxiety ($n=11$), internet gaming disorder ($n=9$), and stress ($n=8$) was widely investigated after verifying the assumption of normality (Shapiro-Wilk's test: $p=0.26$, 0.15 , and 0.24 , respectively). The BSMAS association with anxiety showed medium effect size ($r=0.32$) and high heterogeneity ($I^2=96.85\%$, Cochran's $Q=302.53$, $df=10$, $p<0.001$) with statistically significant results when sample size was tested as a moderator ($F[1,9]=2.35$, $p<0.001$). Similarly, the BSMAS association with depression showed medium effect size ($r=0.32$) and high heterogeneity ($I^2=96.23\%$, Cochran's $Q=518.86$, $df=10$, $p<0.001$), with statistically significant results when sample size was tested as a moderator ($F[1,9]=214.73$, $p<0.001$). Moreover, the BSMAS association with internet gaming disorder showed a moderate-high effect size ($r=0.45$) and high heterogeneity ($I^2=99.69\%$, Cochran's $Q=3597.05$, $df=8$, $p<0.001$) with significant results when sample size was tested as a moderator ($F[1,7]=2535.18$, $p<0.001$). Finally, the BSMAS association with stress showed a moderate effect size ($r=0.34$) and high heterogeneity ($I^2=96.56\%$, Cochran's $Q=289.70$, $df=7$, $p<0.001$),

and was statistically significant when sample size was tested as a moderator ($F[1,6]=147.86$, $p<0.001$).

However, since they have not been thoroughly examined, a qualitative evaluation was undertaken. Negative correlations were found with self-esteem (ranging from -0.18 to -0.29; Dadiotis et al., 2021; Stănculescu, 2023; Žmavc et al., 2022), satisfaction with life ($r=-0.23$; El Abiddine et al., 2024), happiness ($r=-0.11$; Stănculescu, 2023), positive mental health and sense of control ($r=-0.17$ and -0.34 ; Brailovskaia, & Margraf, 2022), conscientiousness, emotional stability, and general subjective well-being ($r=-0.14$, -0.23 , and -0.15 , respectively; Balcerowska et al., 2020). On the other hand, preliminary convergent validity results showed positive association with *Facebook* addiction ($r=0.76$ and 0.72 ; Balcerowska et al., 2020; Copez-Lonzoy et al. 2023), fear of missing out ($r=0.64$; Copez-Lonzoy et al. 2023), loneliness ($r=0.12$ and 0.40 ; Dadiotis et al. 2022; Naher et al. 2022), intensity of social media use ($r=0.54$; Stănculescu, 2023), social anxiety ($r=0.21$; Stănculescu, 2023), need to belong ($r=0.23$; Stănculescu, 2023), anxious attachment ($r=0.43$; Stănculescu, 2023), extraversion ($r=0.07$; Balcerowska et al., 2020), social media engagement and addiction ($r=0.34$; Copez-Lonzoy et al. 2023), nomophobia ($r=0.58$; Pramukti et al. 2023), satisfaction with life obsessive-compulsive disorder ($r=0.33$; Andreassen et al., 2016), and attention-deficit/hyperactivity ($r=.41$; Andreassen et al., 2016). Lastly, only one study (Watson et al., 2020) examined the association between the BSMAS compared to other scales related to SMA such as Social Media Addiction Scale (Al-Menayes, 2015) and Social Media Disorder Scale (van den Eijnden et al., 2016). These findings showed significant associations ($r=0.54$ and $r=0.44$, respectively; Watson et al. 2020).

Criterion validity. The most frequent criteria used to study the criterion validity of the BSMAS's were time spent on the social media platform ($n=9$) and time spent using the smartphone ($n=4$). In most studies, participants were asked to report their daily use, but

sometimes the time-frame was different and not comparable. In sum, the positive correlations were found between BSMAS and daily (i) social media use (ranging from 0.09 to 0.58; Brailovskaia & Margraf, 2022; Chen et al., 2020 a, 2020b; Dadiotis et al., 2021; Lin et al., 2017; Naher et al., 2022; Yam et al., 2019; Žmavc et al., 2022), and (ii) smartphone use (ranging from 0.1 to 0.2; Chen et al., 2020a, 2020b; Yam et al., 2019). Only two studies reported the relationship with smartphone weekly use ($r=0.16$; Leung et al., 2020), and social media weekly use ($r=0.27$ and 0.1 , respectively; Leung et al., 2020; Ruckwongpatr et al., 2024).

Publication bias in BSMAS studies (RQ6)

On a graphical level, a funnel plot was performed for Cronbach's alpha (Figure 7a) and for anxiety, depression, internet gaming disorder, and stress (Figure 8) with Fisher's z-transformed correlation coefficient used to estimate publication bias due to homogeneous data. The graphs showed symmetry and no results deviated significantly from the average effect size. This finding was also confirmed by Begg's test (Cronbach's alpha: Kendall's $\tau=-0.09$, $p=0.76$; depression: Kendall's $\tau=-0.31$, $p=0.02$; anxiety: Kendall's $\tau=-0.09$, $p=0.76$; depression: Kendall's $\tau=-0.13$, $p=0.65$; internet gaming disorder: Kendall's $\tau=-0.33$, $p=0.26$; stress: Kendall's $\tau=-0.07$, $p=0.90$) and Egger's test (Cronbach's alpha: $z=1.81$, $p=0.07$; anxiety: $z=1.81$, $p=0.07$; depression: $z=-0.98$, $p=0.32$; internet gaming disorder: $z=-0.71$, $p=0.47$; stress: $z=0.69$, $p=0.49$). Moreover, the estimation of publication bias was adjusted using the trim-and-fill method. The adjusted funnel plots are shown in Figure 7b for Cronbach's alpha, and in Figure 9 for construct validity measures. More specifically, the trim-and-fill adjusted effect size was confirmed for the overall Cronbach's alpha (adjusted Cronbach's $\alpha=0.33$), and for the construct validity measures with minimal differences (anxiety=0.33, depression=0.33, internet gaming disorder=0.49, stress=0.32).

Discussion

The present study aimed to meta-analyze the psychometric properties of the BSMAS, which is a widely used tool for assessing social media addiction (SMA). The systematic review identified 23 studies supporting its psychometric properties among 17 languages. Therefore, the present study's findings significantly improve the current body of knowledge on the BSMAS. More specifically, the study is the first meta-analysis to simultaneously take into account dimensionality, internal consistency, test-retest reliability, and validity of all previously published studies on this topic. The COSMIN checklist suggested that the internal consistency, responsiveness, and construct validity of the BSMAS showed very good results based on the studies included in the meta-analysis. Although the association with other constructs has been analyzed, measurement invariance across different socio-demographic (e.g., gender, age, etc.) and cultural groups has not been well studied. Only a few studies have analyzed such aspects, and among individuals from different groups, making the results too inconsistent for further meta-analyses. Moreover, criterion validity was investigated in 11 studies. Future research is needed to differentiate between clinical and non-clinical populations due to the lack of shared diagnostic criteria. Further research is also needed for a gold standard in social media addiction diagnoses. Likewise, some psychometric features, such as content validity and measurement error, have yet to be conducted.

In relation to internal consistency, the findings showed that the Cronbach's alpha coefficient of the BSMAS across studies was good (0.84 in the random effect model). The findings overwhelmingly corroborated the unidimensionality of the six-item version of the BSMAS, mostly by performing CFA. However, other methods such as latent class analysis and item response theory have also supported the psychometric properties of the BSMAS. Although a few studies have used these methods, there are also other novel and innovative methodological analyses that could be carried out in future studies to assess the dimensionality of the BSMAS (e.g., bi-factor modeling). Moreover, the moderators examined were insufficient

to account for the heterogeneity for reliability results. For instance, less than half of the included studies reported participants' time spent on smartphones or social media, and sometimes this was reported in various time-frames, such as daily or weekly. Furthermore, assessment setting (i.e., online vs. offline), population type (i.e., university students vs. others), and sample size were not statistically significant moderators. This may have been because the possible categorizations were unbalanced. For example, there were 14 studies that collected data online with only six studies that collected the data offline. Although four moderators were tested, the source of heterogeneity remained unexplained. Further research should conduct other moderators analyses, for example, using cultural dimensions, socioeconomic factors, or psychometric differences in translation processes as moderators. The data available in the published papers did not provide sufficient information to test these hypotheses.

Although the preliminary results were encouraging, test-retest reliability was only investigated in three studies and at different time intervals. The currently available validity results demonstrated medium effect sizes ($r > 0.30$) for four relevant measures, namely anxiety, depression, stress, and internet gaming disorder. In all cases, the high heterogeneity was explained by using sample size as a moderator. Although many studies are conducted in line with available resources, future research should include large samples for more accurate validation. Other associations with related constructs (e.g., loneliness, fear of missing out, nomophobia, etc. were also examined, but the few results available did not allow for an overall evaluation. Moreover, time spent on smartphones or social media was shown to be an effective criterion for BSMAS criterion validity in only a few studies. Therefore, further research is needed to confirm the initial findings.

The results also showed low risk of publication bias. Additionally, these results were also confirmed after trim-and-fill adjustment. Indeed, the overall effect size variations were negligible. Although a first inspection of the funnel plots might suggest that studies with

unsatisfactory psychometric characteristics could not be published (e.g., Cronbach's $\alpha < 0.70$), these results support the good psychometric properties of the BSMAS.

Finally, the present review also provided empirical evidence to support the use of the BSMAS to assess social media addiction. Given the current lack of a gold standard in psychiatric manuals such as the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013), the BSMAS could be used for theoretical reflection, clinical research and practice, and for potential clinical diagnosis and intervention within broader diagnostic frameworks as a valid tool with optimal psychometric properties across countries and settings. Luo et al. (2021) recently proposed a clinical cut-off for the BSMAS, but evidence is still limited and restricted to specific populations. Further research should particularly include clinical samples to achieve this goal. As suggested by the authors of the original version (i.e., Andreassen et al., 2016), who adapted the BSMAS from the BFAS (Andreassen et al., 2012), the BSMAS can be used to assess any addiction related to social media use, in contrast to other tools that focus on a specific social media platform such as Facebook, Instagram and YouTube. This is an advantage considering how quickly new social media platforms can emerge, and how quickly the popularity of some social media platforms can wane. However, other psychometric instruments to assess social media addiction are currently available and have been compared. For example, Watson et al. (2019) compared the BSMAS with the Social Media Disorder Scale (van den Eijnden et al., 2016) and the Social Media Addiction Scale (Al-Menayes, 2015), emphasizing the ability of the BSMAS to capture gender-specific differences in social media addiction. However, there are no psychometric reviews that have compared the currently available instruments. Although the present findings suggest the BSMAS is a candidate as a gold standard for the social media addiction assessment due to its good psychometric properties and established association with the well-being of users (Duradoni et al., 2020), future research should psychometrically compare other instruments to

provide strong recommendations for clinical and research practice. According to Sigerson and Cheng's (2018) review of psychometric instruments, there are several promising scales to assess social media addiction, but ongoing psychometric research is crucial to provide a foundation of valid measurement in this timely research topic.

Limitations

The findings should be considered in light of some of the limitations. Firstly, the quality assessment of the included studies highlighted the extensive use of convenience samples in most studies, particularly university students, which reduces the generalizability of the results. This limitation affects the generalizability of the results. Further psychometric evaluation of the BSMAS should comprise representative samples including older adults, clinical groups, and individuals from different cultural and educational backgrounds so that the psychometric properties of the BSMAS can be examined and compared among individuals from a wide variety of populations. In line with this concern, some studies used relatively small samples for their analysis ($N < 500$). Additionally, the online or offline assessment setting could be a significant source of bias due to poor standardization of procedures.

Notably, although the use of social media platforms to collect the data can be a limitation in terms of it being a convenience sampling procedure, such a way for collecting data can be an advantage for the specific purposes of the BSMAS because selecting participants directly from online platforms allows for the confirmation of an essential selection criterion, namely that the participants are social media users. Moreover, other limitations related to the search strategy, which did not include studies published in language other than English, such as the Turkish BSMAS which was published in its native language (i.e., Demirci, 2019), as well as a relatively small number of other studies, which may introduce selection bias. However, the highest quality journals and consequently the most important papers tend to be published in English.

Future reviews carried out by large multi-lingual research teams could include psychometric research published in other languages. Findings from such studies may impact the findings of the present review by including from populations that are often little studied or with specific cultural peculiarities (e.g., poor or developing countries or countries where there is a significant digital divide).

Moreover, only three studies addressed test-retest reliability at varying time intervals. Therefore, the current evidence makes it difficult to support the stability of the results regarding the temporal reliability of the BSMAS. Further research should include longitudinal studies with standardized intervals to overcome this limitation. Likewise, the validity of BSMAS with respect to external criteria (e.g., time spent on social media, detrimental effects of excessive use) have been inconsistently operationalized across the published studies. Therefore, future studies should implement more standardized approaches to measuring these criteria because such inconsistencies could potentially affect the robustness of the psychometric findings. The currently available findings are so inconsistent that a comprehensive estimation of a coefficient was not possible. Globally, other studies are needed to investigate the psychometric properties of the BSMAS.

Lastly, the primary participants in the studies were mostly university students, and only one study (Rouleau et al., 2023) included psychiatric patients without including more specific information about their mental health. None of the studies included patients with or without a SMA diagnosis. This is evidently related to the fact that there is currently no consensus or gold standard for diagnosing SMA, such as in the fifth edition of the DSM -5 (American Psychiatric Association, 2013). As a consequence, the BSMAS's invariance across different groups, such as age, gender, or other target populations, should be enhanced in further research.

Conclusion

The present systematic review and meta-analysis highlighted encouraging psychometric properties of the BSMAS and its applications for further research. Evidence from 17 different languages confirmed its factor structure invariance and optimal internal consistency. Validity results supported the appropriateness and effectiveness of the inferences made using the instrument. In the near future, more evidence-based, randomized studies targeting various populations/subgroups are needed. Finally, the results could support the reflection on the definition of the nosographic criteria of SMA. In fact, due to the lack of a gold standard for quantifying and categorizing SMA, the use of a standardized method for SMA assessment would aid clinical research and practice in understanding individuals' suffering from this type of addiction.

References

- Al-Menayes , J. (2015). Social media use, engagement and addiction as predictors of academic performance. *International Journal of Psychological Studies*, 7, 86-94.
<https://doi.org/10.5539/ijps.v7n4p86>
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.
- Andreassen, C. S., Billieux, J., Griffiths, M. D., Kuss, D. J., Demetrovics, Z., Mazzoni, E., & Pallesen, S. (2016). The relationship between addictive use of social media and video games and symptoms of psychiatric disorders: A large-scale cross-sectional study. *Psychology of Addictive Behaviors*, 30(2), 252-262.
<https://doi.org/10.1037/adb0000160>
- Andreassen, C. S., & Pallesen, S. (2014). Social network site addiction - an overview. *Current Pharmaceutical Design*, 20(25), 4053-4061.
<https://doi.org/10.2174/13816128113199990616>

- Andreassen, C. S., Torsheim, T., Brunborg, G. S., & Pallesen, S. (2012). Development of a Facebook Addiction Scale. *Psychological Reports*, 110(2), 501-517. <https://doi.org/10.2466/02.09.18.PR0.110.2.501-517>
- Bányai, F., Zsila, A., Király, O., Maraz, A., Elekes, Z., Griffiths, M. D., Andreassen, C.S., & Demetrovics, Z. (2017). Problematic social media use: Results from a large-scale nationally representative adolescent sample. *PloS One*, 12(1), e016983. <https://doi.org/10.1371/journal.pone.0169839>
- Brailovskaia, J., & Margraf, J. (2022). Addictive social media use during Covid-19 outbreak: Validation of the Bergen Social Media Addiction Scale (BSMAS) and investigation of protective factors in nine countries. *Current Psychology*, 43, 13022–13040. <https://doi.org/10.1007/s12144-022-03182-z>
- Brailovskaia, J., Schillack, H., & Margraf, J. (2020). Tell me why are you using social media (SM)! Relationship between reasons for use of SM, SM flow, daily stress, depression, anxiety, and addictive SM use – An exploratory investigation of young adults in Germany. *Computers in Human Behavior*, 113, 106511. <https://doi.org/https://doi.org/10.1016/j.chb.2020.106511>
- Brailovskaia, J., Truskauskaite-Kuneviciene, I., Kazlauskas, E., & Margraf, J. (2023). The patterns of problematic social media use (SMU) and their relationship with online flow, life satisfaction, depression, anxiety and stress symptoms in Lithuania and in Germany. *Current Psychology*, 42(5), 3713-3724. <https://doi.org/10.1007/s12144-021-01711-w>
- Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 50(4), 1088-1101. <https://doi.org/10.2307/2533446>
- Chen, I. H., Ahorsu, D. K., Pakpour, A. H., Griffiths, M. D., Lin, C. Y., & Chen, C. Y. (2020). Psychometric properties of three simplified Chinese online-related addictive Behavior

- instruments among mainland Chinese primary school students. *Frontiers in Psychiatry*, 11, 875. <https://doi.org/10.3389/fpsyt.2020.00875>
- Chen, I. H., Strong, C., Lin, Y. C., Tsai, M. C., Leung, H., Lin, C. Y., Pakpour, A. H., & Griffiths, M. D. (2020). Time invariance of three ultra-brief internet-related instruments: Smartphone Application-Based Addiction Scale (SABAS), Bergen Social Media Addiction Scale (BSMAS), and the nine-item Internet Gaming Disorder Scale-Short Form (IGDS-SF9) (Study Part B). *Addictive Behaviors*, 101, 105960. <https://doi.org/10.1016/j.addbeh.2019.04.018>
- Cohen, J. (1988). *Statistical power analysis for the behavioral Sciences*. Routledge.
- Copez-Lonzoy, A., Vallejos-Flores, M., Capa-Luque, W., Salas-Blas, E., Montero Doig, A. M., Dias, P. C., & Bazo-Alvarez, J. C. (2023). Adaptation of the Bergen Social Media Addiction Scale (BSMAS) in Spanish. *Acta Psychologica*, 241, 104072. <https://doi.org/https://doi.org/10.1016/j.actpsy.2023.104072>
- Dadiotis, A., Bacopoulou, F., Kokka, I., Vlachakis, D., Chrousos, G. P., Darviri, C., & Roussos, P. (2021). Validation of the Greek version of the Bergen Social Media Addiction Scale in Undergraduate Students. *EMBnet J*, 26, e975. <https://doi.org/10.14806/ej.26.1.975>
- Demirci, İ. (2019). Bergen Sosyal Medya Bağımlılığı Ölçeğinin Türkçeye uyarlanması, depresyon ve anksiyete belirtileriyle ilişkisinin değerlendirilmesi. *Anadolu Psikiyatri Dergisi*, 20, 15-22.
- Duradoni, M., Innocenti, F., & Guazzini, A. (2020). Well-being and social media: A systematic review of Bergen addiction scales. *Future Internet*, 12(2), 24. <https://doi.org/10.3390/fi12020024>
- Duval, S. J., & Tweedie, R. L. (2000a). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463. <https://doi.org/10.1111/j.0006-341x.2000.00455.x>

- Duval, S. J., & Tweedie, R. L. (2000b). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, 95(449), 89–98. <https://doi.org/10.1080/01621459.2000.10473905>
- Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109), 629-634. <https://doi.org/10.1136/bmj.315.7109.629>
- El Abiddine, F. Z. E., Aljaberi, M. A., Alduais, A., Lin, C.-Y., Vally, Z., & D. Griffiths, M. (2024). The Psychometric Properties of the Arabic Bergen Social Media Addiction Scale. *International Journal of Mental Health and Addiction*. <https://doi.org/10.1007/s11469-024-01297-x>
- Gomez, R., Zarate, D., Brown, T., Hein, K., & Stavropoulos, V. (2024). The Bergen–Social Media Addiction Scale (BSMAS): Longitudinal measurement invariance across a two-year interval. *Clinical Psychologist*, 28(2), 185-194. <https://doi.org/10.1080/13284207.2024.2341816>
- Griffiths, M. (2005). A ‘components’ model of addiction within a biopsychosocial framework. *Journal of Substance Use*, 10(4), 191-197. <https://doi.org/10.1080/14659890500114359>
- Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., & Welch, V. (2023). Cochrane Handbook for Systematic Reviews of Interventions version 6.4. In. Cochrane. www.training.cochrane.org/handbook
- Islam, M. S., Jahan, I., Dewan, M. A. A., Pontes, H. M., Koly, K. N., Sikder, M. T., & Rahman, M. (2022). Psychometric properties of three online-related addictive behavior instruments among Bangladeshi school-going adolescents. *PloS One*, 17(12), e0279062. <https://doi.org/10.1371/journal.pone.0279062>

- Leung, H., Pakpour, A. H., Strong, C., Lin, Y. C., Tsai, M. C., Griffiths, M. D., Lin, Y.C., & Chen, I. H. (2020). Measurement invariance across young adults from Hong Kong and Taiwan among three internet-related addiction scales: Bergen Social Media Addiction Scale (BSMAS), Smartphone Application-Based Addiction Scale (SABAS), and Internet Gaming Disorder Scale-Short Form (IGDS-SF9) (Study Part A). *Addictive Behaviors*, 101, 105969. <https://doi.org/10.1016/j.addbeh.2019.04.027>
- Lin, C. Y., Broström, A., Nilsen, P., Griffiths, M. D., & Pakpour, A. H. (2017). Psychometric validation of the Persian Bergen Social Media Addiction Scale using classic test theory and Rasch models. *Journal of Behavioral Addictions*, 6(4), 620-629. <https://doi.org/10.1556/2006.6.2017.071>
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications, Inc.
- Luo, T., Qin, L., Cheng, L., Wang, S., Zhu, Z., Xu, J., Chen, H., Liu, Q., Hu, M., Tong, J., Hao, W., Wei, B., & Liao, Y. (2021). Determination the cut-off point for the Bergen social media addiction (BSMAS): Diagnostic contribution of the six criteria of the components model of addiction for social media disorder. *Journal of Behavioral Addictions*, 10(2), 281–290. <https://doi.org/10.1556/2006.2021.00025>
- McGuinness, L. A., & Higgins, J. P. T. (2021). Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods*, 12(1), 55-61. <https://doi.org/10.1002/jrsm.1411>
- Myung, S. K. (2023). How to review and assess a systematic review and meta-analysis article: A methodological study (secondary publication). *Journal of Educational Evaluation for Health Professions*, 20, 24. <https://doi.org/10.3352/jeehp.2023.20.24>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments:

- an international Delphi study. *Quality of Life Research*, 19(4), 539–549.
<https://doi.org/10.1007/s11136-010-9606-8>
- Monacis, L., de Palo, V., Griffiths, M. D., & Sinatra, M. (2017). Social networking addiction, attachment style, and validation of the Italian version of the Bergen Social Media Addiction Scale. *Journal of Behavioral Addictions*, 6(2), 178-186.
<https://doi.org/10.1556/2006.6.2017.023>
- Naher, L., Hiramoni, F. A., Alam, N., & Ahmed, O. (2022). Psychometric assessment of the Bangla version of the Bergen Social Media Addiction Scale. *Heliyon*, 8(7), e09929.
<https://doi.org/10.1016/j.heliyon.2022.e09929>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J., M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Bmj*, 372, n71.
<https://doi.org/10.1136/bmj.n71>
- Pramukti, I., Nurmala, I., Nadhiroh, S. R., Tung, S. E. H., Gan, W. Y., Siaw, Y. L., Yang, Y. N., Griffiths, M. D., Lin, C.J., & Pakpour, A. H. (2023). Problematic Use of Internet Among Indonesia University Students: Psychometric Evaluation of Bergen Social Media Addiction Scale and Internet Gaming Disorder Scale-Short Form. *Psychiatry Investigation*, 20(12), 1103-1111. <https://doi.org/10.30773/pi.2022.0304>
- R Core Team (2021). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rouleau, R. D., Beaugard, C., & Beaudry, V. (2023). A rise in social media use in adolescents during the COVID-19 pandemic: The French validation of the Bergen Social Media

- Addiction Scale in a Canadian cohort. *BMC Psychology*, 11(1), 92.
<https://doi.org/10.1186/s40359-023-01141-2>
- Ruckwongpatr, K., Paratthakonkun, C., Sangtongdee, U., Pramukti, I., Nurmala, I., Angkasith, K., Thanachaisakul, W, Ketchatturat, J., Griffiths, M. D., Kao, K., & Lin, C.Y. (2024). Validity, reliability, and measurement invariance of the Thai Smartphone Application-Based Addiction Scale and Bergen Social Media Addiction Scale. *International Journal of Mental Health Promotion*, 26(4), 293-302.
<https://doi.org/https://doi.org/10.32604/ijmh.2024.047023>
- Schellingerhout, J. M., Verhagen, A. P., Heymans, M. W., Koes, B. W., de Vet, H. C., & Terwee, C. B. (2012). Measurement properties of disease-specific questionnaires in patients with neck pain: A systematic review. *Quality of Life Research*, 21(4), 659–670.
<https://doi.org/10.1007/s11136-011-9965-9>
- Shin, N. Y. (2022). Psychometric properties of the Bergen Social Media Addiction Scale in Korean Young Adults. *Psychiatry Investigation*, 19(5), 356-361.
<https://doi.org/10.30773/pi.2021.0294>
- Shi, L., & Lin, L. (2019). The trim-and-fill method for publication bias: practical guidelines and recommendations based on a large database of meta-analyses. *Medicine*, 98(23), e15987. <https://doi.org/10.1097/MD.00000000000015987>
- Sigerson, L., & Cheng, C. (2018). Scales for measuring user engagement with social network sites: A systematic review of psychometric properties. *Computers in Human Behavior*, 83, 87-105. <https://doi.org/10.1016/j.chb.2018.01.023>
- Stănculescu, E. (2023). The Bergen Social Media Addiction Scale Validity in a Romanian sample using item response theory and network analysis. *International Journal of Mental Health and Addiction*, 21, 2475–2492. <https://doi.org/10.1007/s11469-021-00732-7>

- Sun, Y., & Zhang, Y. (2021). A review of theories and models applied in studies of social media addiction and implications for future research. *Addictive Behaviors*, 114, 106699. <https://doi.org/https://doi.org/10.1016/j.addbeh.2020.106699>
- Szczygieł, K., & Podwalski, P. (2020). Comorbidity of social media addiction and other mental disorders – An overview. *Archives of Psychiatry and Psychotherapy*, 22(4), 7-11. <https://doi.org/10.12740/APP/122487>
- The EndNote Team (2013). *EndNote*. In (Version EndNote X9) [64 bit]. Clarivate.
- Tung, S. E. H., Gan, W. Y., Chen, J. S., Ruckwongpatr, K., Pramukti, I., Nadhiroh, S. R., Chang; Y. L., Lin, C. C., Pakpour, A. H., Lin, C. Y., & Griffiths, M. D. (2022). Internet-related instruments (Bergen Social Media Addiction Scale, Smartphone Application-Based Addiction Scale, Internet Gaming Disorder Scale-Short Form, and Nomophobia Questionnaire) and their associations with distress among Malaysian university students. *Healthcare*, 10(8), 1448. <https://doi.org/10.3390/healthcare10081448>
- van den Eijnden, R. J. J. M., Lemmens, J. S., & Valkenburg, P. M. (2016). The Social Media Disorder Scale. *Computers in Human Behavior*, 61, 478-487. <https://doi.org/10.1016/j.chb.2016.03.038>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1-48. <https://doi.org/10.18637/jss.v036.i03>
- Wang, Z., Yang, H., & Elhai, J. D. (2022). Are there gender differences in comorbidity symptoms networks of problematic social media use, anxiety and depression symptoms? Evidence from network analysis. *Personality and Individual Differences*, 195, 111705. <https://doi.org/https://doi.org/10.1016/j.paid.2022.111705>
- Watson, J. C., Prosek, E. A., & Giordano, A. L. (2020). Investigating psychometric properties of social media addiction measures among adolescents. *Journal of Counseling & Development*, 98(4), 458-466. <https://doi.org/https://doi.org/10.1002/jcad.12347>

- We Are Social, & Meltwater (2024). *Digital 2024 Global Overview Report*. Retrieved August 21, 2024. From: <https://datareportal.com/reports/digital-2024-global-overview-report>"<https://datareportal.com/reports/digital-2024-global-overview-report>
- Yam, C. W., Pakpour, A. H., Griffiths, M. D., Yau, W. Y., Lo, C. M., Ng, J. M. T., Lin, C. Y., & Leung, H. (2019). Psychometric Testing of Three Chinese Online-Related Addictive Behavior Instruments among Hong Kong University Students. *Psychiatric Quarterly*, 90(1), 117-128. <https://doi.org/10.1007/s11126-018-9610-7>
- Yue, H., Zhang, X., Cheng, X., Liu, B., & Bao, H. (2022). Measurement invariance of the Bergen Social Media Addiction Scale across genders. *Frontiers in Psychology*, 13, 879259. <https://doi.org/10.3389/fpsyg.2022.879259>
- Zarate, D., Hobson, B. A., March, E., Griffiths, M. D., & Stavropoulos, V. (2023). Psychometric properties of the Bergen Social Media Addiction Scale: An analysis using item response theory. *Addictive Behaviors Reports*, 17, 100473. <https://doi.org/10.1016/j.abrep.2022.100473>
- Žmavc, M., Šorgo, A., Gabrovec, B., Crnkovič, N., Cesar, K., & Selak, Š. (2022). The protective role of resilience in the development of social media addiction in tertiary students and psychometric properties of the Slovenian Bergen Social Media Addiction Scale (BSMAS). *International Journal of Environmental Research and Public Health*, 19(20), 13178. <https://doi.org/10.3390/ijerph192013178>

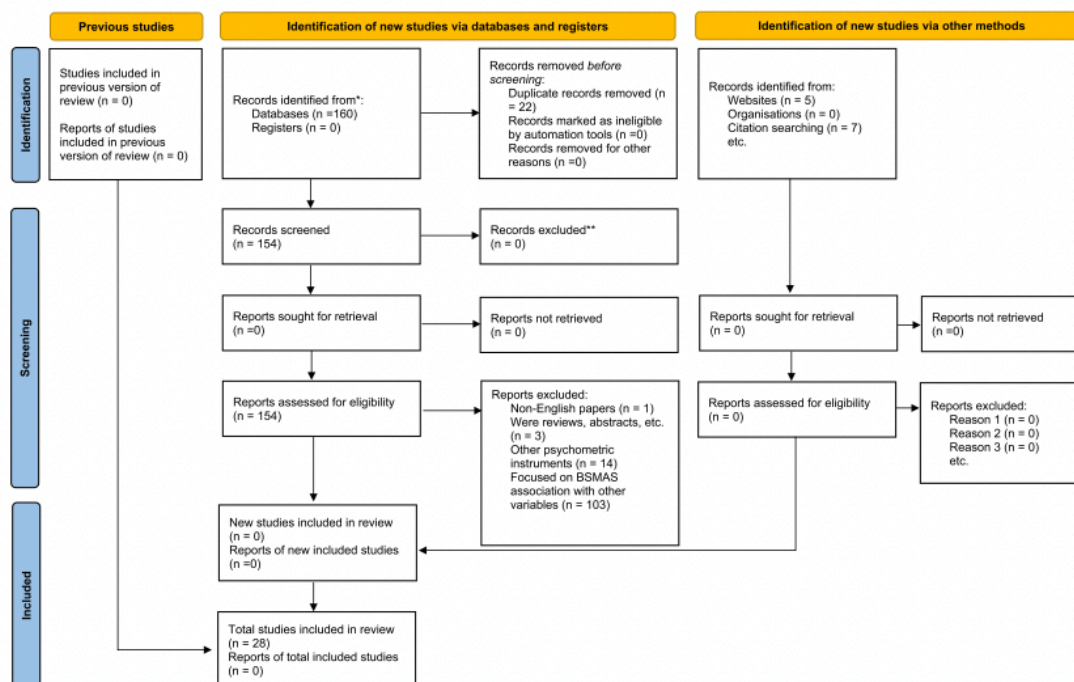


Fig.1 PRISMA 2020 flow diagram the search strategy of the Bergen Social Media Addiction Scale (BSMAS) meta-analysis

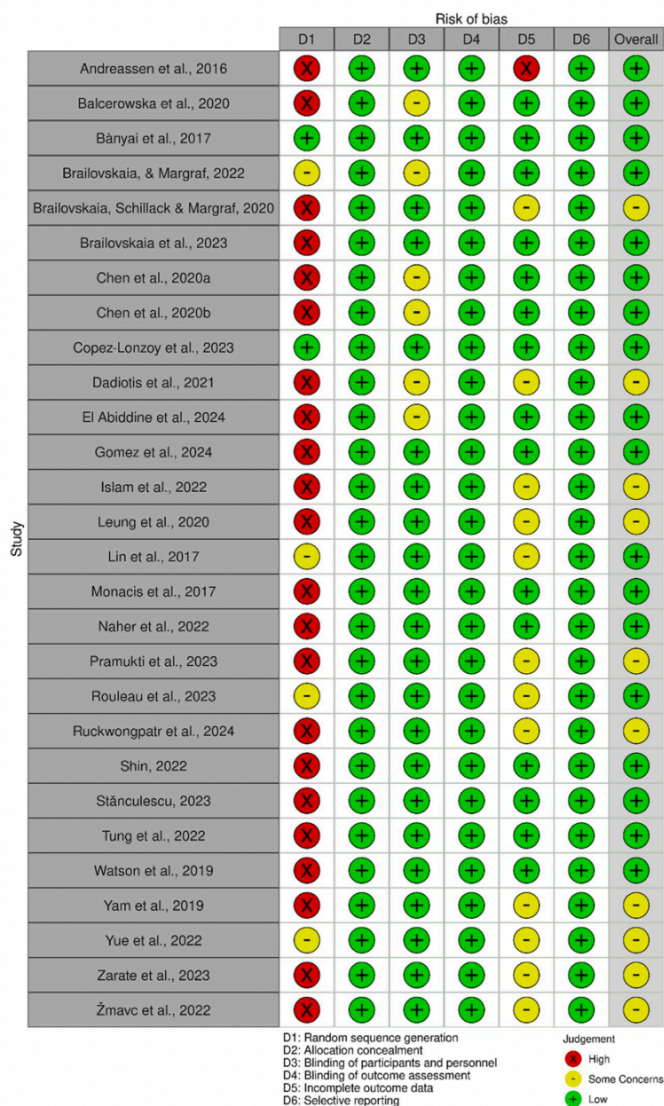


Fig. 2 Traffic-light plot of risk of bias assessment in the studies included in the meta-analysis

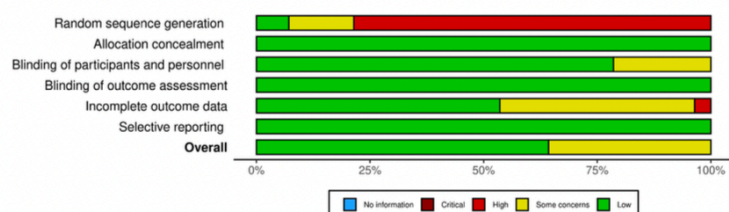


Fig. 3 Summary plot of risk of bias assessment

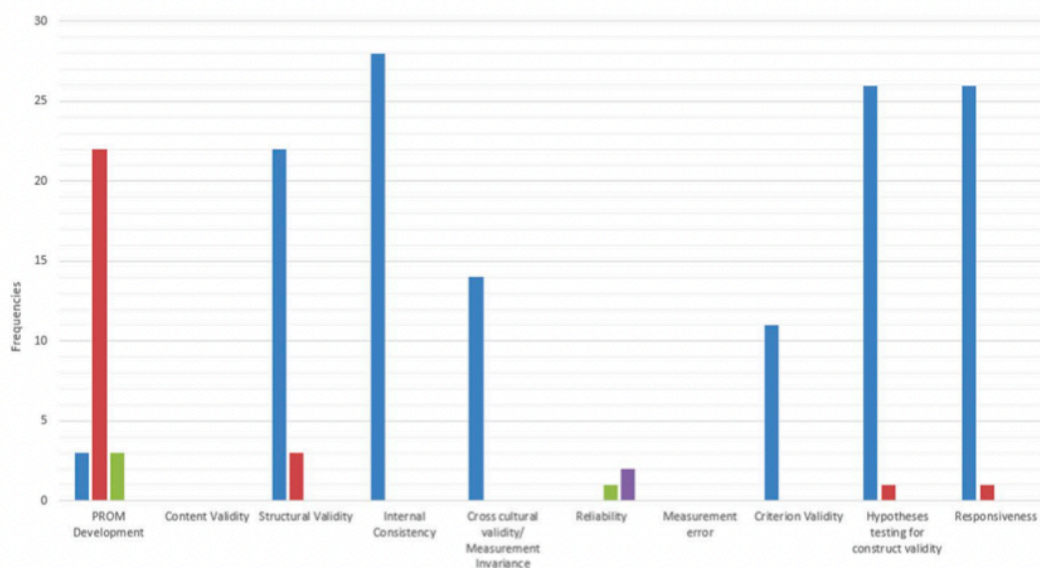


Fig. 4 Summarized methodological quality of each study per measurement property and questionnaire. *Note:* The “not applicable” label was excluded in the current representation; PROM=Patient-reported Outcome Measurement

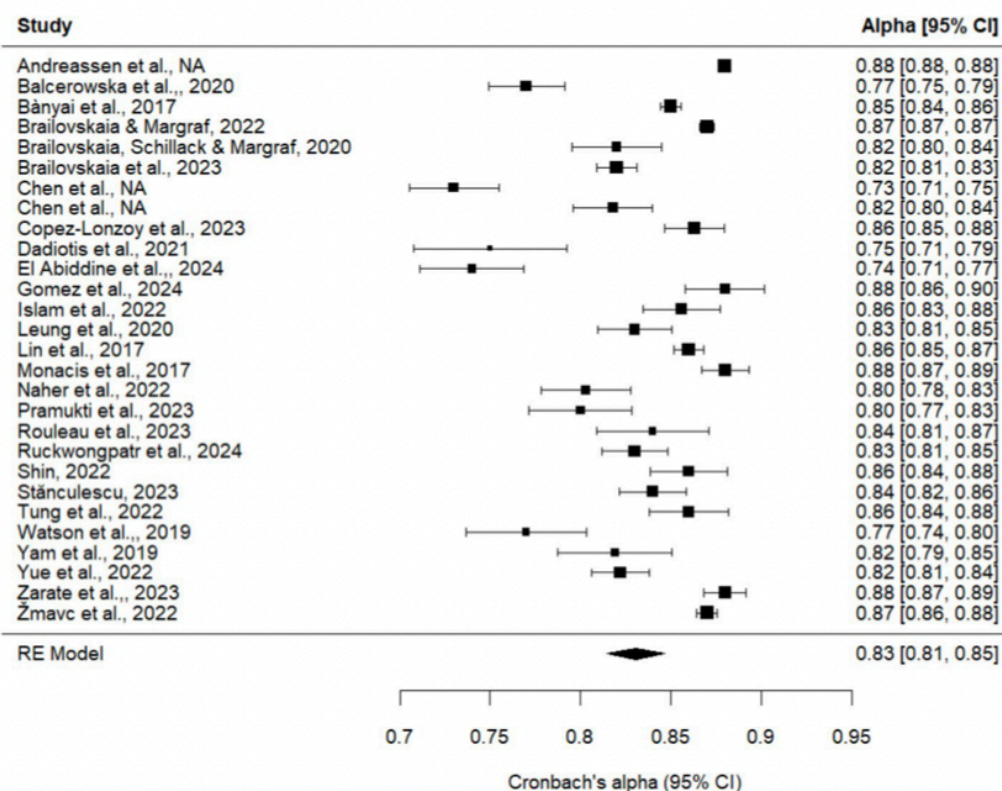


Fig. 5 Forest plot of overall Cronbach's alpha. *Note:* RE Model=random effect model; *original scale development study

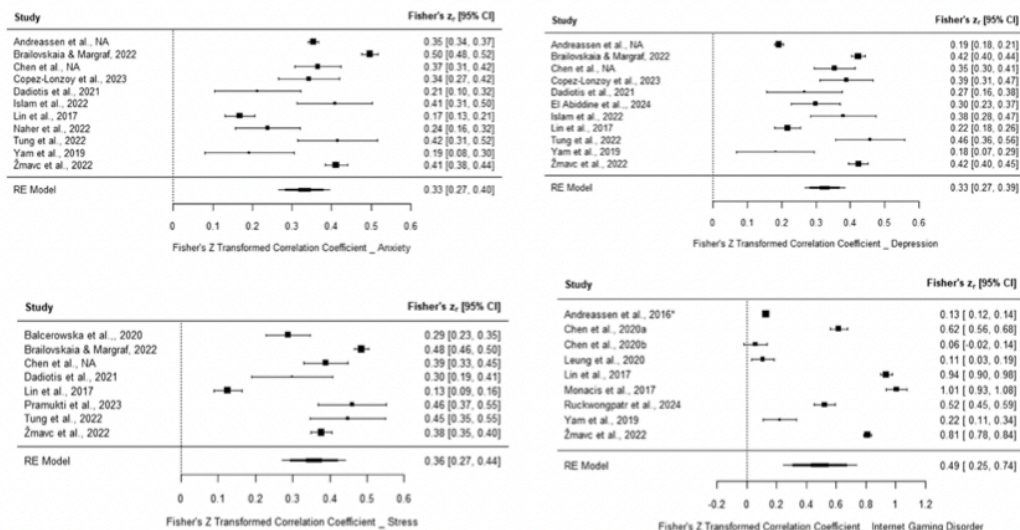


Fig. 6 Forest plot of overall effect size of construct validity measures. *Note:* Re Model=random effect model; * original study

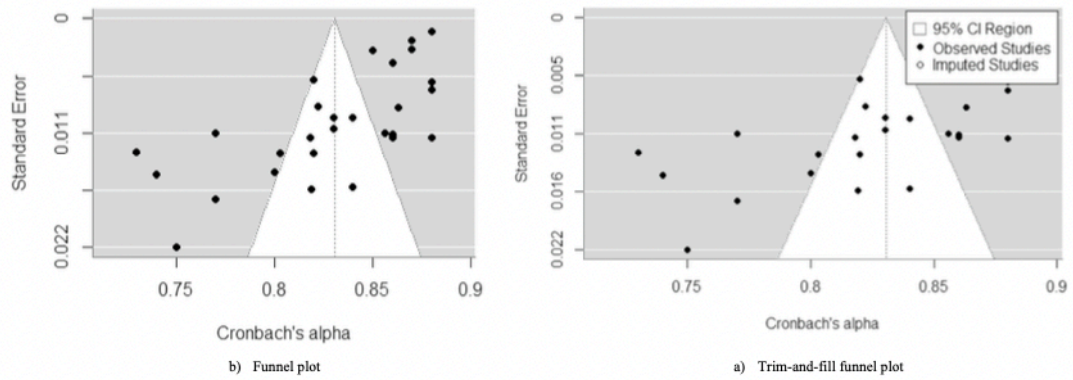


Fig. 7 Funnel plot of included studies' Cronbach's alpha

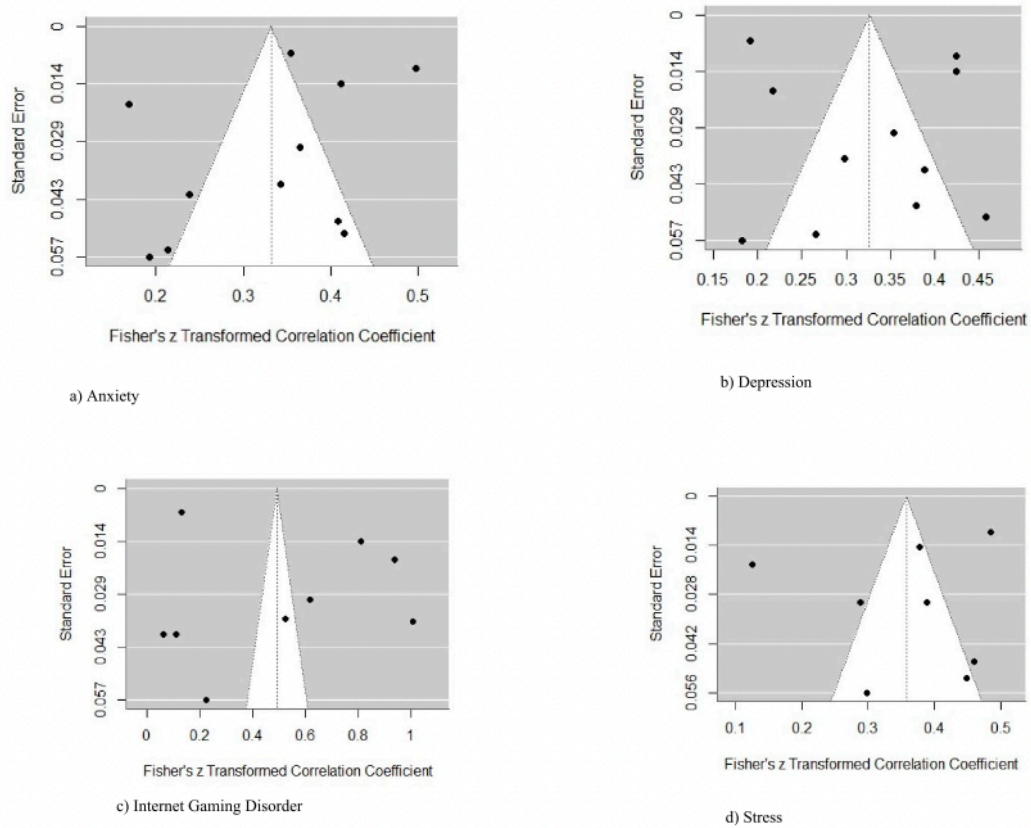


Fig. 8 Funnel plot of overall effect size of construct validity measures

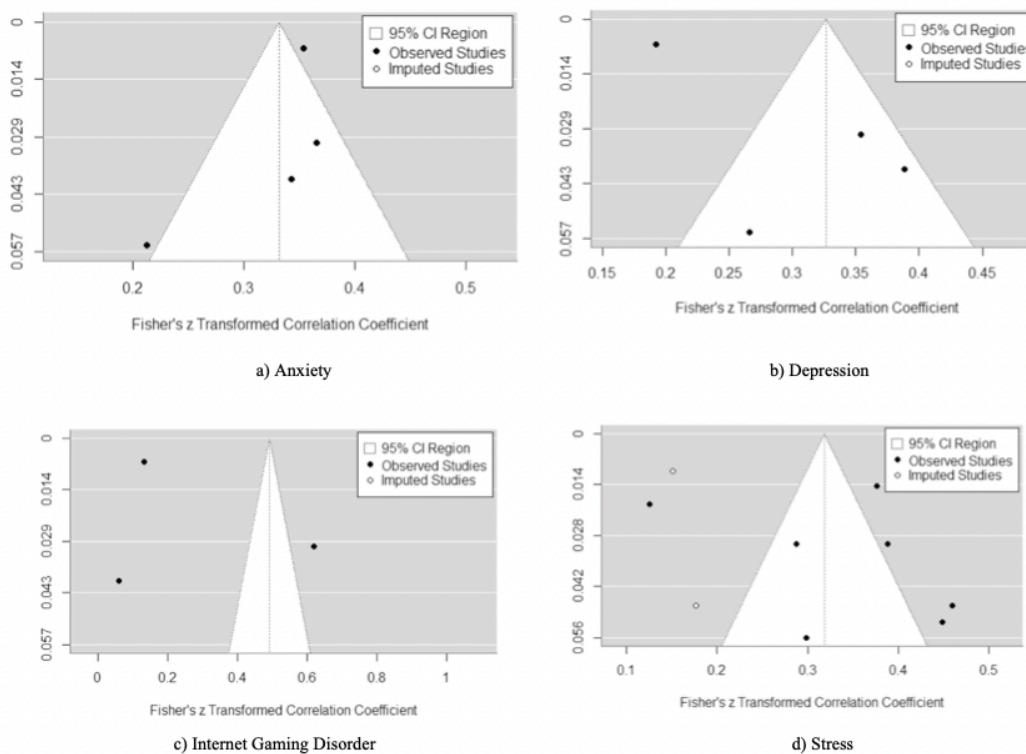


Fig. 9 Trim-and-fill adjusted funnel plot of overall effect size of construct validity measures

Table 1 Studies characteristics ($N=28$)

Study	Language	N (% male)	Mean age $\pm SD$	Target	Assessment setting	Method	Mean BSMAS $\pm SD$	Data analysis	Reported goodness-of-fit indices	Factor loadings range	Cronbach's alpha
Andreasen et al. (2016)*	Norwegian	23,533 (35%)	35.8 $\pm$ 13.3	Adolescents and adults	Online	Cross-sectional	10.3 $\pm$ 4.77	NS	NS	NS	0.88
Balcerowska et al. (2020)	Polish	1099 (28.1%)	21.44 $\pm$ 2.85	Young adults	Online and offline	Cross-sectional	14.76 $\pm$ 4.71	CFA	$\chi^2=251.74$ CFI=0.92, TLI=0.87, RMSEA=0.16 [0.14 0.17]	0.48–0.80	0.77
Bányai et al. (2017)	Hungarian	5961 (49.71%)	16.6 $\pm$ 0.94	Adolescents	Offline	Cross-sectional	NS	CFA + latent class analysis	$\chi^2=5836.19$ CFI=0.95, TLI=0.91, RMSEA=0.07 [0.07 0.08]	0.60–0.81	0.85
Brailovskaia et al. (2022)	Chinese, France, German, Swedish, Polish, Russian, English, and Spanish	9418	NS	Adults	Online	Cross-sectional	NS	PCA	NS	NS	0.81 (China) 0.86 (France) 0.90 (Germany, Sweden) 0.89 (Poland) 0.87 (Russia, United Kingdom, United States) 0.85 (Spain)

Table 1 (continued)

Study	Language	N (% male)	Mean age $\pm$ SD	Target	Assessment setting	Method	Mean BSMAS $\pm$ SD	Data analysis	Reported goodness-of-fit indices	Factor loadings range	Cronbach's alpha
Brailovskaia, Schillack & Margraf (2020)	German	485 (22.3%)	24.75 $\pm$ 6.24	Adults	Online	Cross-sectional	NS	Point-biserial correlations and logistic regression analysis	NS	NS	0.82
Brailovskaia et al. (2023)	Lithuanian and German	2367 (43%)	20.28 $\pm$ 2.03	University students	Online and offline	Cross-sectional	12.89 $\pm$ 5.02	CFA + Latent class analysis	$\chi^2=28.62$ CFI=0.99, TLI=0.99, RMSEA=0.03 [0.02 0.05]	NS	0.81 (Lithuania) 0.83 (Germany)
Chen et al. (2020a)	Chinese	1108 (48.3%)	10.37 $\pm$ 0.95	Children	Offline	Cross-sectional	NS	CFA	$\chi^2=23.93$ CFI=0.99, TLI=0.99, RMSEA=0.04	0.59–0.73	0.73
Chen et al. (2020b)	Chinese	640 (41.56%)	22.34 $\pm$ 3.07	University students	Online	Longitudinal	NS	Time invariance	NS	NS	0.82
Copez-Lonzo et al. (2023)	Spanish	650 (46.5%)	21.4 $\pm$ 2.7	University students	Online	Cross-sectional	12.78 $\pm$ 4.77	CFA	$\chi^2=23.93$ CFI=0.99, TLI=0.99, RMSEA=0.06	0.68–0.82	0.86
Dadiotis et al. (2021)	Greek	325 (18.2%)	21.6 $\pm$ 5.26	University students	NS	Cross-sectional	14.4 $\pm$ 4.37	CFA	NS	NS	0.75

Table 1 (continued)

Study	Language	N (% male)	Mean age $\pm$ SD	Target	Assessment setting	Method	Mean BSMAS $\pm$ SD	Data analysis	Reported goodness-of-fit indices	Factor loadings range	Cronbach's alpha
El Abiddine et al. (2024)	Arabic	757 (35.93%)	21.41 $\pm$ 2.87	University students	Offline	Cross-sectional	NS	CFA + Rash Analysis	CFI=0.97, TLI=0.94	0.40–0.67	0.74
Gomez et al. (2024)	English	276 (71%)	31.86 $\pm$ 9.94	Adults	Online	Longitudinal	20.79 $\pm$ 0.31	CFA + time invariance	$\chi^2=23.95$ CFI=0.96, TLI=0.93, RMSEA=0.08 [0.04 0.11]	0.66–0.80	0.88
Islam et al. (2022)	Bangladeshi	428 (90.89%)	16.13 $\pm$ 1.85	Adolescents	Offline	Cross-sectional	NS	CFA	$\chi^2=29.92$ CFI=0.97, TLI=0.95, RMSEA=0.08 [0.05 0.12]	0.56–0.76	0.86
Leung et al. (2020)	Chinese	642 (41.43%)	22.29 $\pm$ 3.14	Adults	Online	Cross-sectional	NS	CFA	(Hong Kong) $\chi^2=11.24$ CFI=0.99, TLI=0.99, RMSEA=0.03 [0.00 0.08] (Taiwan) $\chi^2=26.89$ CFI=0.97, TLI=0.94, RMSEA=0.08 [0.06 0.11]	0.61–0.72 (Hong Kong) 0.54–0.57 (Taiwan)	0.85 (Hong Kong) 0.82 (Taiwan)
Lin et al. (2017)	Persian	2676 (56.5%)	15.54 $\pm$ 1.21	Adolescents	NS	Cross-sectional	15.24 $\pm$ 4.83	CFA + Rash Analysis	CFI=0.99, TLI=0.99, RMSEA=0.06	0.64–0.83	0.86

Table 1 (continued)

Study	Language	N (% male)	Mean age $\pm$ SD	Target	Assessment setting	Method	Mean BSMAS $\pm$ SD	Data analysis	Reported goodness-of-fit indices	Factor loadings range	Cronbach's alpha
Monacis et al. (2017)	Italian	769 (54%)	21.63 $\pm$ 3.95	Adolescents and young adults	Offline	Cross-sectional	14.2 $\pm$ 5.89	CFA + invariance across gender and age	χ^2 = 66.85 CFI = 0.98, RMSEA = 0.09 [0.07 0.11]	0.46–0.86	0.88
Naher et al. (2022)	Bangladeshi	577 (61.9%)	20.95 $\pm$ 1.92	University student	Offline	Cross-sectional	20.95 $\pm$ 1.92	EFA + CFA + IRT + Network Analysis	χ^2 = 6.62 CFI = 1.00, TLI = 1.00, RMSEA = 0.00	0.57–0.68	0.80
Pramukti et al. (2023)	Indonesian	458 (26%)	22.5 $\pm$ 8.07	University student	Online	Cross-sectional	16.68 $\pm$ 4.50	CFA	CFI = 1.00, TLI = 1.00, RMSEA = 0.00	0.54–0.78	0.80
Rouleau et al. (2023)	French	247 (23.9%)	14.76 $\pm$ 1.6	Adolescents	Online	Cross-sectional	NS	CFA	CFI = 0.98, TLI = 0.96, RMSEA = 0.07	0.56–0.80	0.84
Ruckwongpatr et al. (2024)	Thai	801 (33.1%)	20.69 $\pm$ 3.08	University students	Online	Cross-sectional	15.22 $\pm$ 4.69	CFA + invariance across gender	χ^2 = 19.13 CFI = 0.99, TLI = 0.99, RMSEA = 0.04 [0.01 0.06]	0.56–0.76	0.83
Shin (2022)	Korean	401 (27.2%)	21.9 $\pm$ 1.8	University students	Online	Longitudinal	13.88 $\pm$ 5.36	EFA + CFA	χ^2 = 19.76 CFI = 0.96, TLI = 0.97, RMSEA = 0.08 [0.03 0.12]	0.61–0.77	0.86

Table 1 (continued)

Study	Language	N (% male)	Mean age $\pm$ SD	Target	Assessment setting	Method	Mean BSMAS $\pm$ SD	Data analysis	Reported goodness-of-fit indices	Factor loadings range	Cronbach's alpha
Stănculescu (2023)	Romanian	705 (39%)	30.24 $\pm$ 9.15	Adults	Online	Cross-sectional	11.17 $\pm$ 5.47	CFA + IRT + Network Analysis	CFI = 0.98, TLI = 0.99, RMSEA = 0.06 [0.05 0.09]	0.74–0.94	0.84
Tung et al. (2022)	Malaysian	380 (23.9%)	24.04 $\pm$ 5.07	University students	Online	Cross-sectional	16.8 $\pm$ 5.40	CFA + Rash Analysis	χ^2 = 0.11 CFI = 0.99, TLI = 0.99, RMSEA = 0.04	0.68–0.76	0.86
Watson et al. (2019)	English	440 (49.1%)	17.3 $\pm$ 1.67	Adolescents	NS	Cross-sectional	16.02 $\pm$ U	CFA	χ^2 = 135.39 CFI = 0.82, TLI = 0.70, RMSEA = 0.18 [0.15 0.18]	0.45–0.73	0.77
Yam et al. (2019)	Chinese	307 (32.4%)	21.64 $\pm$ 8.11	University students	Online	Cross-sectional	NS	CFA	χ^2 = 17.06 CFI = 0.98, TLI = 0.96, RMSEA = 0.07	NS	0.82
Yue et al. (2022)	Mongolian	1120 (45.8%)	20.89 $\pm$ 1.4	University students	Offline	Cross-sectional	NS	CFA + invariance across gender	χ^2 = 39.02 CFI = 0.99, RMSEA = 0.05 [0.04 0.07]	0.63–0.73	0.82
Zarate et al. (2023)	English	968 (68.5%)	29.5 $\pm$ 9.36	University students	Online	Cross-sectional	NS	CFA + IRT	χ^2 = 13.35 CFI = 0.99, TLI = 0.99, RMSEA = 0.03 [0.00 0.06]	0.68–0.86	0.88

Table 1 (continued)

Study	Language	N (% male)	Mean age $\pm$ SD	Target	Assessment setting	Method	Mean BSMAS $\pm$ SD	Data analysis	Reported goodness-of-fit indices	Factor loadings range	Cronbach's alpha
Žmavc et al. (2022)	Slovenian	4868 (26.8%)	22.9 $\pm$ 3.19	University students	Online	Cross-sectional	13.7 $\pm$ 5.75	PCA + CFA	CFI = 0.99, TLI = 0.99, RMSEA = 0.09	0.66–0.85	0.87

*original study; NS not specified, CFA confirmatory factor analysis, EFA exploratory factor analysis, IRT item response theory, PCA principal component analysis, CFI Comparative Fit Index, TLI Tucker-Lewis Index, RMSEA root mean square error of approximation

Table 2 COSMIN methodological quality of each study per measurement property and questionnaire

Study	PROM development	Content validity	Structural validity	Internal consistency	Cross cultural validation/ measurement invariance	Reliability	Measurement error	Criterion validity	Hypotheses testing for construct validity	Responsiveness
Andreasen et al. (2016)*	Adequate	NA	NA	Very good	Very good	Inadequate	NA	Very good	Adequate	Very good
Balcerowska et al. (2020)	Doubtful	NA	Very good	Very good	NA	NA	NA	NA	Very good	Very good
Bányai et al. (2017)	Very good	NA	Very good	Very good	Very good	NA	NA	NA	Very good	Very good
Brailovskaia et al. (2022)	Very good	NA	Adequate	Very good	NA	NA	NA	NA	Very good	Very good
Brailovskaia, Schillack & Margraf (2020)	Very good	NA	NA	Very good	NA	NA	NA	NA	Very good	Very good
Brailovskaia et al. (2023)	Adequate	NA	Adequate	Very good	Very good	NA	NA	NA	Very good	Very good
Chen et al. (2020a)	Adequate	NA	Adequate	Very good	Very good	NA	NA	Very good	Very good	Very good
Chen et al. (2020b)	Adequate	NA	NA	Very good	Very good	NA	NA	Very good	Very good	Very good
Copez-Lonzo et al. (2023)	Adequate	NA	Very good	Very good	Very good	NA	NA	NA	Very good	Very good
Dadiotis et al. (2021)	Adequate	NA	Very good	Very good	NA	NA	NA	Very good	Very good	Very good
El Abiddine et al. (2024)	Adequate	NA	Very good	Very good	Na	Na	Na	Na	Very good	Very good
Gomez et al. (2024)	Adequate	NA	Very good	Very good	Very good	Doubtful	Doubtful	NA	NA	NA

Table 2 (continued)

Study	PROM development	Content validity	Structural validity	Internal consistency	Cross cultural validation/ measurement invariance	Reliability	Measurement error	Criterion validity	Hypotheses testing for construct validity	Responsiveness
Islam et al. (2022)	Doubtful	NA	Very good	Very good	NA	NA	NA	Very good	Very good	Very good
Leung et al. (2020)	Doubtful	NA	Very good	Very good	Very good	NA	NA	Very good	Very good	Very good
Lin et al. (2017)	Adequate	NA	Very good	Very good	Very good	NA	NA	Very good	Very good	Very good
Monacis et al. (2017)	Adequate	NA	Very good	Very good	Very good	NA	NA	NA	Very good	Very good
Naher et al. (2022)	Adequate	NA	Very good	Very good	Very good	NA	NA	Very good	Very good	Very good
Pramukti et al. (2023)	Adequate	NA	Very good	Very good	NA	NA	NA	NA	Very good	Very good
Rouleau et al. (2023)	Adequate	NA	Very good	Very good	Very good	NA	NA	NA	Very good	Very good
Ruckwongpatr et al. (2024)	Adequate	NA	Very good	Very good	Very good	NA	NA	Very good	Very good	Very good
Shin (2022)	Adequate	NA	Very good	Very good	NA	Inadequate	NA	NA	Very good	Very good
Stănculescu (2023)	Adequate	NA	Very good	Very good	NA	NA	NA	Very good	Very good	Very good
Tung et al. (2022)	Adequate	NA	Very good	Very good	NA	NA	NA	NA	Very good	Very good
Watson et al. (2019)	Adequate	NA	Very good	Very good	NA	NA	NA	NA	Very good	Very good
Yam et al. (2019)	Adequate	NA	Very good	Very good	NA	NA	NA	Very good	Very good	Very good

Table 2 (continued)

Study	PROM development	Content validity	Structural validity	Internal consistency	Cross cultural validation/ measurement invariance	Reliability	Measurement error	Criterion validity	Hypotheses testing for construct validity	Responsiveness
Yue et al. (2022)	Adequate	NA	Very good	Very good	Very good	NA	NA	NA	Very good	Very good
Zarate et al. (2023)	Adequate	NA	Very good	Very good	NA	NA	NA	Very good	Very good	Very good
Žmavc et al. (2022)	Adequate	NA	Very good	Very good	NA	NA	NA	NA	Very good	Very good

NA not applicable, *PROM* Patient-reported Outcome Measurement