



VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

*The Bergen–Social Media Addiction Scale (BSMAS):
longitudinal measurement invariance across a two-
year interval*

This is the Published version of the following publication

Gomez, Rapson, Zarate, Daniel, Brown, Taylor, Hein, Kaiden and Stavropoulos, Vasileios (2024) The Bergen–Social Media Addiction Scale (BSMAS): longitudinal measurement invariance across a two-year interval. *Clinical Psychologist*, 28 (2). pp. 185-194. ISSN 1328-4207

The publisher's official version can be found at
<https://www.tandfonline.com/doi/full/10.1080/13284207.2024.2341816>
Note that access to this version may require subscription.

Downloaded from VU Research Repository <https://vuir.vu.edu.au/48822/>



The Bergen–Social Media Addiction Scale (BSMAS): longitudinal measurement invariance across a two-year interval

Rapson Gomez, Daniel Zarate, Taylor Brown, Kaiden Hein & Vasileios Stavropoulos

To cite this article: Rapson Gomez, Daniel Zarate, Taylor Brown, Kaiden Hein & Vasileios Stavropoulos (2024) The Bergen–Social Media Addiction Scale (BSMAS): longitudinal measurement invariance across a two-year interval, *Clinical Psychologist*, 28:2, 185-194, DOI: [10.1080/13284207.2024.2341816](https://doi.org/10.1080/13284207.2024.2341816)

To link to this article: <https://doi.org/10.1080/13284207.2024.2341816>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 06 May 2024.



[Submit your article to this journal](#)



Article views: 3909



[View related articles](#)



[View Crossmark data](#)

The Bergen–Social Media Addiction Scale (BSMAS): longitudinal measurement invariance across a two-year interval

Rapson Gomez^a, Daniel Zarate^a, Taylor Brown^a, Kaiden Hein^a and Vasileios Stavropoulos^{a,b}

^aApplied Health, School of Health and Biomedical Science, RMIT University, Melbourne, Australia; ^bDepartment of Psychology, National and Kapodistrian University of Athens, Zografou, Greece

ABSTRACT

Objective: This study aims to examine the longitudinal measurement invariance of the Bergen Social Media Addiction Scale (BSMAS) over a two-year interval, addressing a gap in research on its consistency over time.

Method: Confirmatory Factor Analysis with a robust maximum likelihood chi-square estimator was utilised to assess the BSMAS among 276 adults (mean age = 31.86 years; SD = 9.94 years; 71% male) at three time points across two years. This method evaluates the scale's structural consistency and reliability longitudinally.

Results: The analysis supported full measurement invariance (configural, metric, scalar, and error variance) of the BSMAS, indicating stable measurement properties over time. Furthermore, temporal stability and equivalency of the BSMAS total mean scores were confirmed across the three-time points, suggesting consistent measurement of social media addiction.

Conclusions: The findings validate the BSMAS as a reliable instrument for measuring social media addiction over extended periods. Demonstrating its psychometric stability enhances its utility for longitudinal studies, making it a valuable tool for tracking changes in social media addiction behaviours. These results have significant implications for future research and clinical practice, highlighting the BSMAS's applicability in long-term studies.

KEY POINTS

What is already known about this topic:

- (1) The Bergen Social Media Addiction Scale is a widely utilised instrument for measuring social media addiction.
- (2) There is a lack of research on testing for measurement invariance of the Bergen Social Media Addiction Scale.
- (3) Social media addiction can be a persistent issue across an individual's lifespan thus, measurements of social media addiction need to display invariance across a long time.

What this topic adds:

- (1) Findings showed the Bergen Social Media Addiction Scale displayed support for longitudinal measurement invariance.
- (2) The latent factors of the Bergen Social Media Addiction Scale remain consistent across different time points.
- (3) The Bergen Social Media Addiction Scale can be used to monitor the developmental changes of social media addiction symptoms and clinical treatment effects over a reasonable length of time.

ARTICLE HISTORY

Received 25 December 2023
Accepted 2 April 2024

KEYWORDS


Social media addiction;
Bergen Social Media
Addiction Scale; BSMAS;
confirmatory factor analysis;
longitudinal measurement
invariance

Introduction

Social media (SM) has grown immensely over the last decade with 4.95 billion users globally and 21.3 million in Australia alone (Statista, 2023). While SM offers benefits such as fostering connections and facilitating the sharing of ideas, for a proportion of users excessive use can lead to negative outcomes (Asamoah, 2019). To assess the outcomes of excessive social media use, the Bergen Social Media Addiction Scale (BSMAS;

Andreassen et al., 2016) was created. It is one of the most popular measures used to assess addiction to social media sites (SMS) or social media addiction (SMA). Andreassen and Pallesen (2014) have defined SMA as “being overly concerned about SMS, to be driven by a strong motivation to log on to or use SMS, and to devote so much time and effort to SMS that it impairs other social activities, studies/job, interpersonal relationships, and/or psychological health

CONTACT Kaiden Hein  Kaiden.Hein@rmit.edu.au

 Supplemental data for this article can be accessed at <https://doi.org/10.1080/13284207.2024.2341816>.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

and well-being" (p. 4050). SMA has been linked to emotional, relational, health-related, and performance issues, often manifesting as emotional distress, mental health decline, fear of missing out, and disrupted sleep due to excessive screen time (Al-Samarraie et al., 2021; Andreassen, 2015; Huang, 2022). With global estimates for SMA ranging between 13% (using a severe cut-off) and 25% (using a moderate cut-off; Cheng et al., 2021), and considering the significant impact SMA can have on well-being, validating such measures is crucial.

The BSMAS evolved from the Bergen Facebook Addiction Scale (BFAS; Andreassen et al., 2012) by substituting "Facebook" with "social media" throughout its items. This modification broadened the scope to encompass SM platforms like Facebook, Twitter, Instagram, and similar sites. Rooted in Griffiths' (2005) components model of addiction, the BSMAS addresses six components including; salience (preoccupation with SM), mood modification (involvement with SM improves mood), tolerance (increasing amount of involvement with SM is required to be satisfied), withdrawal symptoms (reduction or preclusion from involvement with SM creates restlessness and negative feelings), conflict (involvement with SMA creates conflicts and causes problems for the individual), and relapse (return to old SMA patterns after a period of control or absence). The BSMAS has six items, each corresponding to one of the addiction components.

In the initial scale development and validation study of the BFAS, Andreassen et al. (2016) obtained BFAS ratings from 423 student participants in Norway. Confirmatory factor analysis (CFA) supported a one-factor model. The study also supported internal consistency reliability, three-week test-retest reliability, and convergent and divergent validity. Additionally, the total score was correlated positively with being female and negatively with being younger. Given the similarity of the BFAS and the BSMAS, these findings can be assumed to apply to the BSMAS.

To date, several studies internationally have examined the psychometric properties of the BSMAS (e.g., Bányai et al., 2017; Chen et al., 2020; Lin et al., 2017; Monacis et al., 2017; Shin, 2022; Zarate et al., 2023). All these studies have supported the unidimensional factor for the BSMAS, as well as, its internal consistency reliability, and convergent and divergent validity. Of note, the participants in the study by Zarate et al. (2023) were recruited at Time 1 ($N = 1097$) in the current study. Although past studies have provided good support for the psychometric properties of the BSMAS, except for Chen et al. (2020), there has been little empirical attention to longitudinal measurement invariance.

Longitudinal measurement invariance implies comparable metric and scalar factorial structures at different time points (Leitgöb et al., 2021). Alternatively, weak or no support for longitudinal measurement invariance suggests that the ratings at the different time points cannot be justifiably compared since the scores can be assumed to be confounded by different measurement and scaling properties. Thus, corresponding empirical information on the BSMAS items' measurement invariance is required to compare ratings at different time points. Support for longitudinal measurement invariance is important as this is necessary for accurately tracking the developmental trajectory of SMA symptoms, assessing the effectiveness of clinical treatments over time, and increasing the generalizability of findings based on BSMAS data collected longitudinally.

While past studies have explored the test-retest reliability of the BSMAS, assessing longitudinal invariance offers a distinct analysis. This approach examines whether observed scores consistently represent the same levels of the underlying latent trait over time, in contrast to test-retest reliability, which focuses on the stability of scores across multiple time points through high correlation measures (see American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (U.S.), 2014). Researchers have reported support for test-retest reliability over a two-week (Shin, 2022), three-week (Andreassen et al., 2016), and three-month (Chen et al., 2020) interval. Related to the BSMAS item network stability, a recent study (submitted) reported invariant network structure and global strength over a one-year interval. Global network invariance refers to equivalent network structures across time points, and global strength invariance refers to equivalent node relations across time points. Although these findings support test-retest reliability and network stability across time, neither of these findings are comparable to the different types of invariances examined using the CFA-based latent variable approach which is the focus of the current study.

Limitations of existing studies

The exploration of the BSMAS psychometric properties reveals a notable gap in the literature, particularly regarding its longitudinal measurement invariance. While Chen et al. (2020) successfully demonstrated this invariance over a three-month span, addiction's complex nature – a persistent and escalating behaviour coupled with psychological dependence (Lüscher et al., 2020; Walker, 1989) – suggests that such a brief period may not suffice for thorough clinical assessments. The scarcity of studies

examining the BSMAS's measurement invariance over more extended intervals, such as a year or more, underscores a significant, unaddressed need for deeper empirical investigation into its long-term reliability and validity.

Aim of the study

Despite the demonstrated unidimensionality and psychometric properties of the BSMAS, a significant gap exists in the literature due to the lack of data on its longitudinal measurement invariance over extended periods. This study aims to fill this gap by examining the BSMAS's longitudinal measurement invariance over a two-year interval, involving three time points, among adults from the Australian general community. Given the limitations in existing data, this study employs CFA to assess the longitudinal measurement invariance of the BSMAS ratings at three annual intervals (i.e., 2021, 2022, and 2023). No specific hypotheses were formulated considering that our measurement invariance examination was primarily exploratory.

Method

Participants

The participants were drawn from the general community, constituting a normative online convenience sample. Participants came from English-speaking countries (e.g., Australia, USA, UK, Canada and New Zealand). Regarding ethnicity, 69.2% identified as White/Caucasian, 18.5% identified as Asian, 6.9% identified as Black/African American, 4% Hispanic/Latino and the remaining 1.4% belonging to other ethnicities. The inclusion criteria comprised English-speaking participants who used social media, while individuals under 18 years old were excluded. To ensure a representative sample, social media users were invited to complete our survey. Although this method employed a convenience sample approach by targeting social media users, it provided equal opportunities for community members to participate in our study. Moreover, using a suggested cut-off value of 26 (Zarate et al., 2023), the number of participants exceeding these values was consistent across waves (4 in the first wave and 4 in the last wave), representing 1.5% of our sample. Regarding individuals with usable scores, responses of 968 English-speaking adults were included at time 1, 462 at time 2, and 276 at time 3. In the current study, the attrition rate from time 1 to time 2 was 52.3% [(968–462)/968], and from time 1 to time 3 was 71.5% [(968–276)/968]. To detect attrition bias in the characteristics of the final sample, we used

t-tests for the BSMAS total scale scores in time 1 who responded and did not respond at time 2 and at time 3 (Miller & Wright, 1995; Mitchell et al., 2022). Supplementary Table S1 shows the descriptives for these variables and the results of the *t*-test. As shown, respondents and nonrespondents differed in the BSMAS total score for both time points. However, as the effect sizes for both these differences were small, it was interpreted that attribution bias had little effect on the scores collected in the study.

Only the 276 participants who completed ratings at all time points were involved in this study. Soper's (2022) software for computing sample size requirements for CFA models was used to evaluate the sample size requirement for the present study. For this, the anticipated effect size was set at 0.3, power at 0.8, the number of latent variables at 3 (covering the three time points), the number of observed variables at 18 (covering the three time points), and probability at .05. The analysis recommended a minimum sample size of 200. Our sample size ($N = 276$) exceeds this recommendation. Further details on the formula can be found on Soper's (2022) website.

Measures

All data was collected online. At the start of the study (time 1), participants provided demographic information, including age, gender, ethnicity, highest education level completed, employment status, and relationship status. They also completed ratings of the BSMAS (Andreassen et al., 2016) at three different time intervals, one year apart (in 2021, 2022, and 2023).

The BSMAS has six items with a time reference of the past year. An example item is: "Do you spent a lot of time thinking about social media and planning the use of social media?" Items are responded to on a five-point Likert scale ranging from very rarely (1) to very often (5). Therefore, higher symptom scores indicate higher symptom severity. The internal reliability for the BSMAS instrument was very good in the present study (Cronbach $\alpha = .88, .9, \text{ and } .9$; McDonald's $\omega = .88, .9, \text{ and } .91$ for time 1, time 2, and time 3, respectively).

Procedure

Upon approval from the Victoria University Human Research Ethics Committee (HRE20-169), the study was advertised using nonelectronic (i.e., word of mouth) and electronic (i.e., email, social media) methods. Time 1 data was collected between August 2019 and August 2020. Participants were invited to register for the study via a Qualtrics link available on social media (i.e., Facebook,

Instagram, Twitter), the Federation University websites and digital forums (i.e., reddit.com). Individuals informed through nonelectronic methods provided an email address to which the Qualtrics link could be sent. The link took them to the Plain Language Information Statement (PLIS). Those wishing to participate were directed to click a button to agree to informed consent. This was followed by questions seeking sociodemographic information and a number of questionnaires. Only the BSMAS is of relevance to the present study. Participants completed the online survey using a computer at their chosen location. At the end of completing the steps at time 1, participants were requested to voluntarily provide their email address to be included in prospective data collection wave(s) and sign the study consent form digitally (box ticking). Twelve months later (between August 2021 and August 2022), those who consented received follow-up emails requesting their voluntary participation in the survey. This included an identical survey link (i.e., PLIS, email provision for the second wave, consent form and survey questions). In all, 462 participated in the second data collection wave. A comparable procedure between August 2022 and August 2023 was used for collecting wave 3 data. The study's inclusion criteria included being an adult (i.e., over 18 years old) and engaging in any form of online activity, such as online gaming. The exclusion criteria were straightforward, encompassing only incomplete and invalid responses. Due to the inclusion of questionnaires addressing one's level of distress, those who had a current untreated severe mental illness were instructed (also included in the plain language information statement) not to participate to avoid any unforeseen/indirect emotional impact. Beyond these conditions, no further inclusion or exclusion criteria were applied to ensure a broad and inclusive participant pool.

To ensure the highest standards of data integrity and reliability, a detailed quality control strategy was meticulously executed throughout the data lifecycle. This involved a thorough mapping of each step in the data workflow, from initial collection through to the final dataset preparation, with a keen focus on ensuring the reversibility of actions to protect raw data integrity. Modifications to data were carefully documented under distinct filenames, incorporating versioning to facilitate clear traceability. Additionally, our processes were standardized and explicitly documented to enable replication by future researchers, ensuring consistent and reliable results. Predefined data structures and collection templates were employed to further enhance data consistency. In an effort to foster transparency and facilitate reproducibility, the dataset has been made accessible alongside this submission.

Statistical procedures

All the CFA models were conducted using *Mplus* (Version 7) software (Muthén & Muthén, 2012). Although the BSMAS scores are ordinal, they can be treated as continuous, as there are five response options (Rhemtulla et al., 2012). The robust maximum likelihood chi-square (MLR) estimator was used for all analyses. The $MLR\chi^2$ was utilized to examine the goodness-of-fit of the CFA models. Similar to other χ^2 values, large sample sizes lead to $MLR\chi^2$ values being exaggerated. As well as providing the $MLR\chi^2$, *Mplus* also provides approximate or practical fit indices, including the root mean squared error of approximation (RMSEA), the comparative fit index (CFI), the Tucker-Lewis Index (TLI), and the Standardized Root Mean Square Residual (SRMR). The current study supports the fit of the one-factor model at all three time points. Model fit was evaluated using the root mean squared error of approximation (RMSEA), comparative fit index (CFI), Tucker-Lewis index (TLI), and the standardized root mean square residual. According to Hu and Bentler (1998), RMSEA, values $< .06$ = good fit, < 0.08 = acceptable fit, and > 0.08 to $.10$ = marginal fit. For CFI and TLI, values $\geq .95$ = good fit, and $\geq .90$ = acceptable fit. For the SRMR, it should be less than $.05$ for a good fit (Hu & Bentler, 1999), although values smaller than $.10$ may be interpreted as acceptable (Schermelleh-Engel et al., 2003). Hu and Bentler (1999) have recommended a two-index approach for evaluating model fit that includes fit in terms of the SRMR value and either the TLI, CFI, or RMSEA. For the current study, a model was considered acceptable if the SRMR value was smaller than $.05$, and either TLI or CFI or RMSEA showed acceptable fit.

Before testing longitudinal measurement invariance, the one-factor BSMAS model fit at time 1, time 2, and time 3 was examined. In these models, the ratings for all six items were loaded onto a single latent factor, with uncorrelated error variances. Additionally, for model identification purposes, the variance of the latent factors was fixed at one.

An extended single-group CFA model that included the ratings at all time points was used for longitudinal measurement invariance. Supplementary Figure 1 shows our path diagram for evaluating longitudinal measurement invariance. The model combines the unidimensional factor models for time 1, time 2, and time 3. However, the models are connected at each time point, with correlated like error variances and latent factors. Using this model, we assessed longitudinal measurement invariance across time points by sequentially comparing models with progressively stricter constraints. This process involved testing for configural invariance, metric

invariance, scalar invariance, and uniqueness invariance. Supplementary Table S2 provides details for the steps involved in the analysis. In brief, the CFA procedures for measuring measurement invariance involve comparing progressively more constrained models that test several levels of invariance. In the context of longitudinal measurement invariance, this involves showing that (i) the latent factor structure remains the same between time points (baseline or configural invariance); (ii) the associations/strengths of like items with their latent factors are the same at different time points (metric or loading invariance); (iii) the item intercepts of like items are the same at different time points (scalar or intercept/threshold invariance); and (iv) the item uniqueness variances of like items are the same at different time points (uniqueness or the unique factor invariance). For comparing the various nested CFA models we used both the χ^2 difference test, and the differences in the approximate fit indices (CFI and RMSEA). For the latter, differences to reject invariance for both factor loadings and thresholds are set at $\Delta\text{CFI} > .01$, and $\Delta\text{RMSEA} > -.015$, respectively (F. F. Chen, 2007).

Results

Demographic information of the sample

Table 1 provides background information on the 276 participants involved in the study. As shown, their age ranged from 18 to 62 years (mean = 31.86 years;

SD = 9.94 years) and included 196 men (71%; mean age = 31.92 years, SD = 10.84 years) and 75 women (32.9%; mean age = 32.12 years, SD = 10.84 years). Additionally, five individuals (1.8%) did not identify their gender. No significant age difference was found across men and women, $t(269) = 0.1496$, $p = 0.882$. Regarding sociodemographic background, slightly more than half the number of participants (66.5%) reported being employed, and most reported having completed at least secondary education (97%). Racially, most of the participants identified themselves as “white” (69.2%), and slightly less than half the number of participants (42.4%) indicated that they were involved in romantic relationships. Based on the gold standard of clinical diagnosis, Luo et al. (2021) have proposed a score of 24 (out of a total scale score of 30) to distinguish between those at risk and not at risk for SMA. Based on this cut-off, the frequencies for those at risk for the total sample at time 3 was 9 or 3.3%.

Preliminary analyses

Initially, the mean and standard deviation scores for the six items of BSMAS were computed. These are displayed in Supplementary Table S3. Following this, the fit of the one-factor BSMAS model was examined at the three-time points. Table 2 displays the results of these analyses. As shown, based on Hu and Bentler's (1998) recommendations, and their two-index approach for

Table 1. Frequencies and descriptive statistics of the sample.

Variables	Frequencies	Means (SD)
Number of participants	276	
Age (Time 1), years		31.86 (9.94); Range 18–62
Sex		
Female	75 (27.2%)	32.12 (10.84)
Male	196 (71%)	31.92 (9.66); $t(269) = 0.149$, $p = 0.882$
Other	5 (1.8%)	
Employed	183 (66.5%)	
Race		
White	191 (69.2%)	
Black/African American	19 (6.9%)	
Asian	51 (18.5%)	
Hispanic/Latino	11 (4.0)	
Others (Aboriginal, Pacific Islander, Mixed)	4 (1.5%)	
Highest educational level		
Primary	3 (1.1%)	
Secondary	71 (25.7%)	
Technical	30 (10.9%)	
Some University	130 (59.1%)	
Others	5 (1.8%)	
Relationship		
Is involved	117 (42.4%)	
Video game		
Video game usage	276 (100.0%)	
Years playing preferred game		21.26 (9.026); Max/Min = 0/43
Hours on video games -weekdays		8.27 (9.342); Max/Min = 0/75
Hours on video games -weekends		6.34 (9.496); Max/Min = 0/32

SD = Standard Deviation; Max/Min = Maximum/Minimum.

Table 2. Fit values for the BSMAS 1-factor CFA model at time 1, time 2 and time 3.

Model at	$\chi^2(df)$	CFI	TLI	RMSEA (90% CI)	SRMR
Time 1	23.952(9)**	.960	.934	.078 (.041–.116)	.036
Time 2	37.510(9)***	.931	.885	.107 (.073–.144)	.047
Time 3	29.307(9)***	.958	.930	.090 (.055–.128)	.033

CI = confidence interval; RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis Index; SRMR = standardized root mean square residual.

*** $p < .001$; ** $p < .01$.

evaluating model fit (Hu & Bentler, 1999), at time 1, time 2, and time 3, the SRMR values. Overall, therefore, the findings were interpreted as indicating good fit for the one-factor model at all three time points. Notwithstanding this, the findings can be interpreted as indicating sufficient fit for the one-factor BSMAS model at all three-time points. Supplementary Table S4 shows the factor loadings, intercepts, and error variances for the BSMAS one-factor CFA model at time 1, time 2, and time 3.

Longitudinal measurement invariance for the BSMAS 1-factor CFA model across time 1, time 2 and time 3 based on the $\Delta\chi^2$

Table 3 summarizes the results for testing longitudinal measurement invariance for the BSMAS one-factor CFA model. As shown, for the configural invariance model (M1), the CFI and TLI indicate an acceptable fit, and the RSMEA indicated a good fit. Thus, there was support for the configural invariance model. The table also shows that for both differences in chi-square ($\Delta\chi^2$) and approximate fit indices (Δ CFI and Δ RMSEA) there no difference between the full metric invariance model (M2) and the configural invariance model (M1), the full scalar invariance model (M3) and the full metric invariance model

(M2), the full scalar invariance model (M3) and the full error variances invariance model (M4). Thus, there was support for full metric, scalar, and error variance invariance models, respectively, i.e., full longitudinal measurement invariance.

Post hoc analysis of temporal stability of the BSMAS latent factor

Table 4 presents the intercorrelations among the latent factors of the BSMAS at three time points, alongside the results of a one-way repeated measures ANOVA examining differences in total observed scores across these time points. The table illustrates significant correlations indicating strong temporal stability and reports the ANOVA findings that suggest no significant difference in total mean scores over time, $F(2, 274) = 2.775, p = .064$. Also, a 1-way repeated ANOVA showed no difference in the latent mean scores [mean (standard error) for time 1 = 20.786 (.314); time 2 = 10.076 (.386); time 3 = 20.38 (.309)] across the three time points, $F(2, 274) = 2.775, p = .064$.

Table 4 also includes the intraclass corrections between the different time points (Shrout & Fleiss, 1979). As they were all large effect sizes (Cohen, 1988), temporal stability can be assumed. Additionally, Supplementary Figure S3 shows the relevant Bland-

Table 3. Results of the test for longitudinal measurement invariance for the BSMAS 1-factor CFA model across time 1, time 2 and time 3 based on the $\Delta\chi^2$.

#	Model (M)	Model fit values				Models compared	Differences in fit			
		$\chi^2 (df)$	CFI	TLI	RMSEA (90%CI)		df	χ^2	CFI	RMSEA
1	Configural	239.492 (114)	.936	.915	.063 (.052–.074)					
2	Metric – M1 with all loadings free	241.371 (124)	.941	.927	.059 (.047–.070)	M2 – M1	10	4.447	.005	-.004
3	Scalar invariance – M2 with all thresholds free	259.014 (136)	.938	.930	.057 (.047–.068)	M3 – M2	12	15.419	-.003	-.002
4	Error variance invariance – M3 with all error variance free	274.970 (144)	.934	.930	.057 (.047–.068)	M4 – M3	8	15.723	-.004	.000

CI = confidence interval; RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis Index. All MLR χ^2 values were significant ($p < .001$).

Table 4. Temporal stability of the BSMAS latent factor.

Measurement Point	Correlation with Time 1	Correlation with Time 2	Mean (SE)	ANOVA Results
Time 1	–	.527* (.49)	20.786 (.314)	$F(2, 274) = 2.775, p = .064$
Time 2			10.076 (.386)	
Time 3	.657* (.61)	.418* (.39)	20.38 (.309)	

* $p < .001$, indicating significant correlations.

Correlations are reported to illustrate the temporal stability of the latent factor across three measurement points. The ANOVA results are provided to show the comparison of total mean scores over time, indicating no significant difference. The values in parenthesis are the intraclass correlations.

Altman plots. The Bland-Altman plot, is a graph comparing two measurements techniques (Bland & Altman, 1986, 1999), or as in our case the same measurement at two time points. The differences between the two techniques/time points are plotted against the averages of the two techniques. Horizontal lines are drawn at the mean difference, and at the limits of agreement, which are defined as the mean difference plus and minus 1.96 times the standard deviation of the difference. If these limits do not exceed the maximum allowed difference between methods Δ , the two methods are considered to be in agreement and may be used interchangeably. As seen in the plots, this is generally the case for differences at time 1 and time 2, time 1 and time 3, and time 2 and time 3, stability in the scores at all three time points can be assumed.

Discussion

Summary of study findings

The findings of the present study contribute novel insights into the longitudinal measurement invariance of the BSMAS items among adults, examining data across three distinct time points over a two-year period. Given the significant prevalence of SMA, estimated to affect 13% to 25% of individuals globally (Cheng et al., 2021), and its detrimental effects on well-being, including heightened anxiety, depression, and reduced self-esteem (Al-Samarraie et al., 2021; Huang, 2022), validating reliable measures of SMA is of paramount importance. This validation process ensures that assessments of SMA accurately reflect consistent characteristics of addiction over time, facilitating better understanding, monitoring, and intervention strategies for individuals affected by excessive social media use. The findings showed that there was support for configural invariance (same factor structure pattern), full metric (same factor loadings), full scalar (same response level), and full unique factor (same unique variances) invariance. Overall, these findings show strong support for longitudinal measurement invariance in adults for the BSMAS items across three time points spanning two years. Although unrelated to invariance, we also found support for strong temporal stability for the latent factors across the three time points, including equivalency for latent mean scores across the three time points.

Meaning of our invariance findings

In the context of the current study, the support for configural invariance indicates that the same overall factor structure (one factor in the current study) holds across the

three time points. The support for metric invariance indicates that the strength of the associations of the items with the BSMAS latent factors are the same for like items at all three time points. Scalar invariance indicates that individuals will endorse the same level of observed scores for the same latent trait scores at different time points. The support for error variance invariance indicates that the reliabilities of the BSMAS items are the same for like items at all three time points. Studies assessing the test-retest reliability of the BSMAS have reported moderate to strong reliability over short intervals (Bányai et al., 2017; Lin et al., 2017; Monacis et al., 2017; Shin, 2022; Zarate et al., 2023). Our findings indicate that the BSMAS also maintains its reliability over longer periods, extending up to two years. To date, Chen et al. (2020) is the only study that has examined measurement invariance, and their findings corroborate ours, indicating that the BSMAS demonstrates configural, scalar, and error variance invariance. Although that study also supported measurement invariance, it was limited as this was demonstrated for two time points over a three-month interval. Therefore, our findings can be considered new and novel in this area.

Two other findings not directly related to our primary goal of evaluating measurement invariance are worthy of note. Firstly, we found support for strong temporal stability for the latent factors across the three time points. This suggests that SMA has stronger stability over longer intervals than previously reported (Andreassen et al., 2012; Chen et al., 2020; Shin, 2022). Secondly, prior to the measurement invariance analyses, we examined the fit for the one-factor BSMAS model at time points 1, 2, and 3. Overall, consistent with previous studies, we interpreted our findings as showing at least adequate fit for the one-factor model (e.g., Bányai et al., 2017; Chen et al., 2020; Lin et al., 2017; Monacis et al., 2017; Shin, 2022; Zarate et al., 2023).

Clinical and practical implications

The strong support for longitudinal measurement in adults for the BSMAS items across three time points spanning two years suggests that the BSMAS scores are not confounded by biases related to scaling and measurement issues and can therefore be justifiably compared over this interval. Thus, the BSMAS items can be used to monitor the developmental changes of SMA symptoms and accurately monitor clinical treatment effects over a reasonable length of time. However, this recommendation must be viewed with caveats in mind.

Strictly speaking, our findings apply to the SMA symptoms included in the BSMAS and not to SMA symptoms in general. However, there are reasons to suspect that such a possibility cannot be ruled out.

Specifically, as the content of the items in the BSMAS is based on the components model of addiction (Griffiths, 2005) that is thought to capture core addiction symptoms (salience/preoccupation, mood modification, tolerance, withdrawal symptoms, conflict, and relapse), the study's findings could be relevant to other measures of SMA and possibly SMA in general.

Limitations

Although the current study has delivered original and valuable information regarding longitudinal measurement invariance about the BSMAS symptom ratings across two years, the findings and interpretations must be considered with several limitations in mind. First, SMA ratings are influenced by other factors, such as age and gender (Andreassen et al., 2012). Not controlling for these variables in the present study may have confounded findings. Second, as all participants in this study were from the general community and not selected randomly, our findings may be further confounded and limited in terms of generalization, including their application to those with the potential for clinical levels of SMA. Third, all data used were collected using a self-rating questionnaire (i.e., the BSMAS). Again, it is possible that the ratings were influenced by this method and as such our results may be subject to confounding by common method variance. Fourth, our findings have been obtained from a single study, and therefore replication is essential. Fifth, the use of convenience sampling, while strategic for targeting users of social media to assess symptoms via the BSMAS, introduces inherent limitations to the generalizability of our findings. This method may not capture a fully representative cross-section of the broader population, possibly affecting the accuracy of symptom prevalence rates. Future research should consider employing more diverse sampling techniques to mitigate potential biases and enhance the representativeness of study findings. Finally, while we establish sufficient power for the study, it may still be possible that the findings would have been different if the sample had been larger. Given these limitations, future research is needed in this field, controlling for the limitations noted above.

Conclusions

In conclusion, the present study has established the stability of the BSMAS scores across a two-year period, effectively showing that it is not influenced by scaling or measurement inconsistencies. This underscores the BSMAS's utility in monitoring SMA

symptoms and evaluating the effectiveness of interventions over time. Highlighting its novelty, this research is pioneering in investigating the scale's longitudinal measurement invariance for self-reported SMA symptoms over such an extended timeframe. It offers potential substantial contributions to both theoretical and practical realms in SMA, advising clinicians and researchers to leverage these insights for longitudinal studies on SMA symptoms.

Acknowledgments

The authors wish to thank all the individuals who participated in the study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Ethical standards – animal rights

All procedures performed in the study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Ethical approval for the study was granted by the following institutions: Victoria University Human Research Ethics Committee (application no. HRE20-169). This article does not contain any studies with animals performed by any of the authors.

Informed consent

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

Availability of data and materials

The data used in the analysis is available on request from the corresponding author.

Authors' contribution

RG: contributed to the literature review, framework formulation, and the structure and sequence of theoretical arguments. VS: contributed to the framework formulation, the structure and sequence of theoretical arguments, data collection, and reviewed the final form of the manuscript.

DZ: contributed to the framework formulation, the structure of theoretical arguments, data collection, and reviewed the final form of the manuscript.

KH: Contributed to the literature review, reviewed the final form of the manuscript and final submission.

TB: contributed to the literature review, reviewed the final form of the manuscript, data collection.

Funding

VS was supported by the Australian Research Council (Discovery Early Career Researcher Award; DE210101107). RB and TB did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors for this research.

References

- Al-Samarraie, H., Bello, K. A., Alzahrani, A. I., Smith, A. P., & Emele, C. (2021). Young users' social media addiction: Causes, consequences and preventions. *Information Technology & People*, 35(7), 2314–2343. <https://doi.org/10.1108/ITP-11-2020-0753>
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for educational and psychological testing*. AERA.
- Andreassen, C. S. (2015). Online social network site addiction: A comprehensive review. *Current Addiction Reports*, 2(2), 175–184. <https://doi.org/10.1007/s40429-015-0056-9>
- 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.PRO.110.2.501-517>
- Asamoah, M. K. (2019). The two side coin of the online social media: Eradicating the negatives and augmenting the positives. *International Journal of Ethics Education*, 4(1), 3–21. <https://doi.org/10.1007/s40889-018-0062-6>
- Bányai, F., Zsila, Á., 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. *Public Library of Science One*, 12(1), e0169839. <https://doi.org/10.1371/journal.pone.0169839>
- Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 327(8476), 307–310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
- Bland, J. M. & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8(2), 135–160. <https://doi.org/10.1177/096228029900800204>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Cheng, C., Lau, Y. C., Chan, L., & Luk, J. W. (2021). Prevalence of social media addiction across 32 nations: Meta-analysis with subgroup analysis of classification schemes and cultural values. *Addictive Behaviors*, 117, 106845. <https://doi.org/10.1016/j.addbeh.2021.106845>
- 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* (2nd ed.). Erlbaum.
- Griffiths, M. D. (2005). A 'components' model of addiction within a biopsychosocial framework. *Journal of Substance Use*, 10(4), 191–197. <https://doi.org/10.1080/14659890500114359>
- Huang, C. (2022). A meta-analysis of the problematic social media use and mental health. *International Journal of Social Psychiatry*, 68(1), 12–33. <https://doi.org/10.1177/0020764020978434>
- Hu, L. T. & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424. <https://doi.org/10.1037/1082-989X.3.4.424>
- Hu, L. T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Leitgöb, H., Seddig, D., Schmidt, P., Sosu, E., Davidov, E., Cernat, A., & Sakshaug, J. W. (2021). Longitudinal measurement (non) invariance in latent constructs. In A. Cernat & J. W. Sakshaug (Eds.), *Measurement error in longitudinal data* (pp. 211–258). Oxford University Press. <https://doi.org/10.1093/oso/9780198859987.003.0010>
- 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 Addiction*, 6(4), 620–629. <https://doi.org/10.1556/2006.6.2017.071>
- 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>
- Lüscher, C., Robbins, T. W., & Everitt, B. J. (2020). The transition to compulsion in addiction. *Nature Reviews Neuroscience*, 21(5), 247–263. <https://doi.org/10.1038/s41583-020-0289-z>
- Miller, R. B. & Wright, D. W. (1995). Detecting and correcting attrition bias in longitudinal family research. *Journal of Marriage and the Family*, 57(4), 921–929. <https://doi.org/10.2307/353412>
- Mitchell, M. M., Fahmy, C., Clark, K. J., & Pyrooz, D. C. (2022). Non-random study attrition: Assessing correction techniques and the magnitude of bias in a longitudinal study of reentry from prison. *Journal of Quantitative Criminology*, 38(3), 755–790. <https://doi.org/10.1007/s10940-021-09516-7>

- 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>
- Muthén, L. K. & Muthén, B. O. (2012). *Mplus Version 7 user's guide*. Muthén & Muthén.
- Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods, 17*(3), 354. <https://doi.org/10.1037/a0029315>
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research, 8*(2), 23–74.
- 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>
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428. <https://doi.org/10.1037/0033-2909.86.2.420>
- Soper, D. S. (2022). *A-priori sample size calculator for structural equation models*. [Software]. Available from <https://www.danielsoper.com/statcalc>
- Statista. (2023). *Social media usage worldwide*. <https://www.statista.com/study/12393/social-networks-statista-dossier/>
- Walker, M. B. (1989). Some problems with the concept of “gambling addiction”: Should theories of addiction be generalized to include excessive gambling? *Journal of Gambling Behavior, 5*(3), 179–200. <https://doi.org/10.1007/BF01024386>
- 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>