ORIGINAL ARTICLE

Problematic Social Network Use: Structure and Assessment

Covadonga González-Nuevo[1] · Marcelino Cuesta[1] · Álvaro Postigo[1] ·
Álvaro Menéndez-Aller[1] · José Muñiz[2]

Abstract

Using social networks (SNs) inappropriately can lead to psychological problems. The objective of this study was to develop a new measuring instrument of problematic use of SNs. The sample comprised 1003 participants over 18 years old ($M = 42.33$; $SD = 14.32$). Exploratory factor analysis was performed with a randomly selected 30% of the sample, and confirmatory factor analysis with the remaining 70%. The reliability of the instrument was estimated, and evidence of validity in relation to the variables—anxiety, depression and satisfaction with life—was obtained. The new scale demonstrated a two-dimensional structure (GFI =0.99; RMSEA= 0.06), with one factor of negative social comparison ($\alpha = 0.94$) and another of addictive consequences ($\alpha = 0.91$). Clear evidence of validity related to other variables was found. The new scale demonstrated good psychometric properties. The advantage of this questionnaire is that it assesses not only excessive use but also social comparison through SNs.

Keywords Social networks · Problematic social media use · Social comparison · Addictive social media use

Time spent in social networks (SNs) has been linked to various psychological problems such as depression, anxiety and psychological well-being (Frost & Rickwood, 2017; Hussain & Griffiths, 2021). However, in order to clarify this relationship, it is necessary to determine which behaviours in SNs are truly problematic, as certain activities in SNs may not be problematic but positive, such as using SNs as a form of social support (Gilmour et al., 2020).

Various terms have been used to refer to maladaptive SNs use, such as problematic social media use (Bányai et al., 2017), social media disorder (Van Den Eijnden et al., 2016) and Facebook addiction (Andreassen et al., 2012). Despite this broad range of terminology, there are two main approaches for evaluating the problematic use of SNs, each one related to the two existing models for problematic internet use (Matute, 2016). One approach conceives problematic SNs use as a problem of addiction, and

---

✉ Covadonga González-Nuevo
  gonzalezvcovadonga@uniovi.es

1 University of Oviedo, Oviedo, Spain

2 University of Nebrija, Madrid, Spain

thus focuses on symptoms usually present in behavioural addiction (Allahverdi, 2021; Andreassen et al., 2012; Blanca & Bendayan, 2018; Van Den Eijnden et al., 2016). The other approach applies Caplan's (2010) model of problematic internet use to SNs.

Studies based on Caplan's (2010) model not only address typical symptoms of behavioural addiction, such as mood-related and negative consequences, but also include other traits which would denote problematic use (Marino et al., 2016). One such trait is particularly important in problematic SNs use and could be assessed alongside addictive traits and is named negative social comparison (Verduyn et al., 2020). Negative social comparison on SNs, defined as comparisons that make a person feel inferior to others on SNs, has been related to negative mental health consequences (Appel et al., 2015). To the best of our knowledge, no instruments have been developed so far that evaluate this specific behaviour on SNs.

Social comparison was defined by Festinger (1954) as the tendency to use other people as sources of information to examine one's own abilities and ways of behaving, thinking, or feeling. These comparisons, which are innate in humans, are intensified on SNs due to the almost limitless content available (Perloff, 2014). This comparison can have positive or negative consequences, depending on which one compares oneself with. There are two possibilities: either the person compares themselves with someone they consider inferior (downward comparison) and their self-esteem improves or they compare themselves with someone considered superior (upward comparison), which is a common situation on SNs and is known to have negative effects on a person's mental health (De Lenne et al., 2018; De Vries et al., 2018; Luong et al., 2019; Verduyn et al., 2020).

The assessment of social comparison on SNs has mostly depended on unvalidated questionnaires. Two types of questionnaires have been used without the corresponding psychometric analysis. On the one hand, there are unvalidated questionnaires that have been constructed for a specific study, such as the questionnaire by Lee (2014), in which the author makes no mention of factorial structure or reliability. However, it has continued to be used in subsequent studies (De Vries & Kühne, 2015; Schmuck et al., 2019). Another example of an "ad hoc" questionnaire about SNs comparison is in Cramer et al. (2016), in which only reliability was addressed. On the other hand, researchers have modified questionnaires aimed at evaluating social comparison in other contexts to assess social comparison on SNs. For example, a common practice to adapt them to the SNs context is to change the content of the items, mainly, by adding the word "Facebook". For instance, Liu et al. (2017) adapted an instrument aimed at assessing comparison in day-to-day life, the Iowa Netherlands Comparison Measure (Gibbons & Buunk, 1999), to the context of SNs.

The most commonly studied psychological problems related to problematic SNs use are anxiety, depression and psychological well-being, with gambling and cyberbullying as other less frequently related issues (Feijóo et al., 2021). Problematic use, understood as a purely addictive problem, has been associated with higher levels of depression and anxiety (Barbar et al., 2021; Keles et al., 2020; Malaeb et al., 2021; Seabrook et al., 2016; Youssef et al., 2021), stress (Hussain & Griffiths, 2018) and poorer psychological well-being (Marino et al., 2018b). Problematic SN use, when considered a context for negative social comparison, has also been associated with greater levels of depression (Yoon et al., 2019) and anxiety-depressive symptoms (Schmuck et al., 2019), as well as with lower levels of psychological well-being (Arias-de la Torre et al., 2020; Huang, 2017). Recent meta-analyses have also studied problematic SNs use by combining both types of uses, addictive and comparative, which they have found to be both positively related to depression (Vahedi & Zannella, 2021; Yoon et al., 2019).

In summary, there seems to be a relationship between the development of psychological disorders and the problematic use of SNs. The strength of this relationship, however, largely depends on how the use of SNs is defined, and therefore measured. To date, research has addressed this relationship by understanding the problematic use of SNs as a purely addictive problem. Our study aims to bring a broader approach by developing a tool that also assesses social comparison on SNs, which is not only a fundamental characteristic of SNs but it is also completely different to the social comparison that takes place in other contexts. This questionnaire, meant to be used with any SNs, is the first one in the Spanish context. Therefore, our main objective in this study is the development and validation of a measuring instrument that will allow standardized, rigorous measurement of problematic SNs use. In order to do that, we will perform the appropriate psychometric analyses to provide evidence of item quality and relationship with other variables, along with reliability and dimensionality.

Method

Participants

The sample was initially composed of 1,059 participants from the Spanish general population over 18 years old. The final sample was reduced to 1,003 following the removal of 5.29% of the sample for having more than two incorrect answers in the attentional control scale (described in more detail in the "Instruments" section). The participants were aged between 18 and 83 years old ($M = 42.33$; $SD = 14.32$), with 75.5% of them being women and almost two-thirds (64.81%) of the participants having university-level qualifications. Table 1 shows the distribution of gender and level of education both in the sample and the Spanish general population.

Instruments

Problematic Use of SNs (PUS) Questionnaire

In the development of the new instrument, problematic use of SNs (PUS), we followed the criteria established by the European Federation of Psychologists' Associations (EFPA) for test evaluation (Evers et al., 2013) and the standards for Educational and Psychological

Table 1. Distribution of gender and educational attainment in the sample and in the Spanish general population

		% Primary education	% Under university qualifications	% University-level qualifications	% Total
% Spanish population	Men	40.2	23.1	36.7	49
	Women	34.1	23.3	42.7	50.9
	Total	36.8	21.6	41.3	
% Sample	Men	3.3	43.9	52.8	24.5
	Women	2.4	28.9	68.7	75.5
	Total	2.59	32.53	64.81	

Evaluation (AERA, APA, NCME, 2014), as well as the recommendations from the current literature on psychometry (Downing & Haladyna, 2006; Lane et al., 2015; Muñiz, 2018; Muñiz & Fonseca-Pedrero, 2019).

A review of the literature was undertaken to construct a set of items representative of the behaviours that comprise problematic use of SNs (54 items). All of the items were worded directly (Suárez-Álvarez et al., 2018; Vigil-Colet et al., 2020). Problematic SN use was defined in the same way as in Marino et al. (2018a), as the use of SNs that produces negative consequences in people's psychological well-being. From this definition, the items were developed in accordance to the two theoretical areas of problematic SN use: the addiction-related consequences of SNs use and negative social comparison. Addictive use would reflect the interference of SNs use in everyday life, while negative social comparison would reflect the comparisons made on SNs which puts the individual in a situation of inferiority. While these two facets were created to cover the theoretical aspects of maladaptive use of SNs, the internal structure of the test, however, was expected to reflect unidimensionality. This is because both addictive use and negative comparisons on SNs would both be uses of SNs with negative consequences on people's lives. There are also meta-analyses that have examined the problematic use of SNs as a combination of the two uses (Vahedi & Zannella, 2021; Yoon et al., 2019).

A detailed qualitative and quantitative analysis was carried out to determine the representativeness of the content (Sireci & Faulkner-Bond, 2014) with the participation of 21 doctors of psychology. The expert assessment had two parts: (a) rating the clarity of the item wording on a scale from 1 to 10 and (b) rating how strongly each item belonged to each hypothesized dimension of problematic use of SNs. The expert judges were also able to comment on each item. From the data thus obtained, we applied the following three criteria to remove low-quality items: (a) inter-rater agreement below 70% with respect to the measured dimension, (b) the lower limit of the 95% confidence interval of Aiken's V below 0.7 (Charter, 2003) and (c) unfavourable comments. Thus, we removed 34 items, with a final scale containing 20 items. The scores in the final scale for Aiken's V were 0.88 [95% CI 0.82–0.93], indicating an excellent level of agreement about the clarity of the item wording (Penfield and Giacobbi 2004). The final instrument consisted of 20 Likert-type item with five response categories (1 completely disagree, 5 completely agree).

Satisfaction with Life Scale (SWLS; Diener et al., 1985)

This instrument is a scale that measures life satisfaction and it has five items. Participants are asked how much they agree with each statement, responding using a 5-point Likert-type scale (from 1 = completely disagree to 5 = completely agree). Reliability estimated using the α coefficient for the Spanish adaptation was 0.88 (Vázquez et al., 2013). The α coefficient in the present study was 0.82.

Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983)

We used the Spanish adaptation from Terol et al. (2007). This is a questionnaire containing 14 items, with two subscales of 7 items each, responses to which are given on a Likert-type scale from 0/3. One subscale, HADS_A, assesses the level of anxiety, and the other, HADS_D, assesses the level of depression. Higher scores in each subscale indicate greater anxiety and depression, respectively. The internal consistency of the two scales in the

Spanish version was 0.86 (Quintana et al., 2003). The α coefficient in the present study was 0.84 for HADS_A and 0.77 for HADS_D.

Attentional Control Scale

We included an attentional control scale in order to detect participants who responded randomly to the different questionnaires. It was composed of 10 items that asked participants to choose a specific response (e.g., in this question, mark completely agree).

Procedure

We used snowball sampling via SNs to obtain the sample. Participants had to be 18 years old or older to participate. Data collection was done through an online questionnaire, anonymously and voluntarily, with informed consent being given before starting. Both the questionnaire items and the attentional control items were presented to the participants in random order. The participants received no reward for their participation.

Data Analysis

We used the SPSS 24 statistics package (IBM Corp, 2016) to calculate the descriptive statistics, differential item functioning (DIF), Pearson correlations and the canonical correlation. We used FACTOR 10.10.02 (Lorenzo-Seva & Ferrando, 2013) to perform the exploratory factor analysis (EFA) and to calculate reliability coefficients. Finally, we used the Mplus8 program (Muthén & Muthén, 2017) to carry out the confirmatory factor analysis (CFA).

To examine the internal structure of the test, we randomly divided the sample into two subsamples. We carried out an EFA in the first subsample, made up of 30% of the total sample (305 subjects), and a CFA in the second subsample (698 subjects). To check that the data was suitable for EFA, we used KMO and Bartlett's statistic. The analysis was done with a polychoric correlation matrix given the ordinal nature of the variables and the high number of items with kurtosis and skewness values greater than |1| (Ferrando Piera, 2021; Muthen & Kaplan, 1992). The method of estimation was robust unweighted least squares (RULS), following the guidelines in the current literature (Lloret-Segura et al., 2014). We determined the number of factors using the optimal implementation of parallel analysis (PA) procedure (Calderón et al., 2019; Timmerman & Lorenzo-Seva, 2011).

Subsequently, in order to confirm the dimensionality indicated by the EFA, we performed a CFA with the remaining 70% of the sample. The method of estimation used was unweighted least squares estimates with standard errors and a mean- and variance-adjusted chi-square test (ULSMV). The indices of fit were *comparative fit index* (CFI), *Tucker-Lewis index* (TLI), *root mean square residual* (RMSR) and *root mean square error of approximation* (RMSEA). A good fit was observed if CFI >0.95, TLI >0.95, RMSR < 0.08 and RMSEA < 0.05 (Hu & Bentler, 1999).

Following both factor analyses, we carried out an item analysis in accordance to the classical test theory model. To assess the discrimination indices of the items, the corrected item-test correlations were calculated, with values higher than 0.2 considered acceptable (Muñiz et al., 2005; Muñiz & Fonseca-Pedrero, 2019). We also examined whether the items had an impact as a function of the variable sex. In the items that did show an impact, we assessed DIF via the logistic regression procedure (Gómez-Benito et al., 2013).

We used Cronbach's (1951) α coefficient and Mcdonald's (1999) ω coefficient to examine reliability.

We examined evidence of validity in relation to other variables by calculating the Pearson correlation between the new instrument and the following variables: (a) the HADS anxiety and depression scales, and (b) the Satisfaction with Life Scale (SWLS). To observe the overall relationship between problematic SNs use and psychological distress, a canonical correlation was carried out between the two dimensions found in the PUS and mood-related variables (anxiety, depression and life satisfaction). To avoid over-representation of high educational level women, all correlational analyses have been done weighting the sample to resemble to the population structure.

Internal evidence of convergent validity between the two subscales of the PUS instrument was assessed using the average variance extracted (AVE). This was done using the method described in Fornell and Larcker (1981), with values above 0.50 considered satisfactory (Hair et al., 2009). Evidence of discriminant validity between the two scales in the instrument was determined by comparing whether the AVE values for each scale were higher than the square of the correlation between them (Fornell-Larcker criterion).

In order to improve the interpretation of the results of the questionnaire, we constructed norm-referenced scores for each of the two subscales. To determine whether it was necessary to differentiate between the sexes, an independent samples weighted t test was performed.

Results

We first performed an EFA with a subsample of 305 subjects. Prior to this analysis, we confirmed the suitability of the data for EFA via the KMO test (KMO = 0.87) and Bartlett's test ($p \leq 0.001$). An initial EFA was done to confirm whether the data fitted to an essentially unidimensional structure, which gave an unsatisfactory fit (Table 2). Despite the value of GFI and TLI being greater than 0.95, indicating an adequate fit, neither RMSEA nor RMSR gave an acceptable fit, and in addition, the PA suggested the extraction of two factors.

We subsequently performed a second EFA to determine the fit of the two-dimensional structure, which produced an adequate fit with a GFI and TLI that was higher than the first

Table 2. Exploratory factor analysis of the PUS questionnaire

	Unidimensional EFA	Bidimensional EFA[1]	Bidimensional EFA[2]
GFI	0.95	0.99	0.99
TLI	0.95	0.99	0.99
RMSEA	0.10	0.05	0.06
RMSR	0.12	0.06	0.06
Explained variance	48.4%	61.3%	63.68%

Note. n = 305; *GFI* goodness of fit index; *TLI* Tucker-Lewis index; *RMSR* root mean square of residuals; *RMSEA* root mean square error of approximation; unidimensional EFA, single factor model using 20 items; bidimensional EFA[1], two-factor model with 20 items; bidimensional EFA[2], two-factor model using 18 items

EFA and an acceptable value of RMSEA and RMSR (see Table 2). However, we eliminated two items from the test (items 5 and 10) as they loaded low and equally on both factors. They were also removed for theoretical reasons as they assessed similar aspects as other items with higher factor loadings. The test was reduced to 18 items, which were analysed by a third EFA, giving adequate fit supported by the GFI and TLI being greater than 0.95, the explained variance and the RMSEA giving a value between 0.05 and 0.08 as well as the RMSR being lower than 0.08. The fit indices of all three EFAs are shown in Table 2.

The correlation between factors was 0.59. These results indicate that the test can be understood as a bidimensional scale. Based on the distribution of the items in each of the factors, the first factor could be called "negative social comparison" and the second "addictive consequences".

The bidimensional model was tested via a CFA with 698 subjects and 18 items. Table 3 shows the values for CFI and RMSEA indicating a good fit. It also shows the factor loadings, which were very high in both factors. In addition, the table includes all the discrimination indices (DI), which were all higher than 0.56 in the first factor and higher than 0.55 in the second factor. Items 1, 12 and 17 showed impact but none of them exhibited gender DIF.

The first PUS factor, called "negative social comparison", had an alpha coefficient of 0.94 and an omega coefficient of 0.94. The second PUS factor, called "addictive consequences", had an alpha coefficient of 0.91 and an omega coefficient of 0.91.

Convergent validity between scales was acceptable for both the negative social comparison factor (AVE = 0.60) and the addictive consequences factor (AVE = 0.56), with AVE values above 0.5. For discriminant validity, we compared the square of the correlation between the two factors ($r^2_{xy} = 0.26$) with the value of AVE for each factor. As the value of AVE for both factors was greater than the square of the correlation, adequate evidence of discriminant validity between scales was found.

With regard to the relationships with other variables, Table 4 shows the Pearson correlations between problematic use of SNs, as measured using the new PUS instrument, and the variables depression, anxiety and satisfaction. The correlation between the HADS subscales (anxiety and depression) and the PUS subscales (negative comparison and addictive consequences) was statistically significant and positive. In addition, the SWLS (satisfaction) correlated negatively with both PUS subscales.

Similarly, the canonical correlation between the two PUS subscales and the three psychological trait scales (anxiety, depression and satisfaction) was 0.52, and the redundancy coefficient was 0.16 (16% common variance).

Finally, we norm-referenced the scores based on percentiles differentiating between the sexes, as the independent sample weighted t test indicated statistically significant differences. Women had significantly higher scores than men in both the negative social comparison subscale ($p = 0.003$) and the addictive consequences subscale ($p = 0.016$). The norm-referenced scores for negative social comparison and addictive consequences for both men and women are shown in Table 5 and 6, respectively.

Discussion and Conclusions

The objective of our study was to develop and validate the problematic use of SNs (PUS) scale. The PUS is novel because (i) it is the only self-report, as far we know, that not only studies the potential addiction-related consequences of SNs use but also focuses on how

Table 3. Confirmatory factor analysis and discrimination index for the items in the PUS questionnaire

Items	F.L. Social comparison	F.L. Addictive consequences	DI^1	DI^2
1	When I see content from influencers or celebrities, I feel inferior	0.73		0.60	
2	When I see content from my friends or people I know, I feel inferior	0.81		0.74	
3	When I'm not on social networks, I feel an impulse to go online that is hard to resist		0.78		0.68
4	When I publish my content, I worry that it will be made fun of	0.68		0.62	
5	When I see what celebrities or influencers publish, I feel bad about myself	0.80		0.64	
6	When I see what my contacts publish, I feel that they have a better life than me	0.78		0.66	
7	I spend too much time using social networks		0.71		0.67
8	I don't get enough sleep because of using social networks		0.70		0.55
9	I have tried to spend less time on social networks, but I have not been successful		0.72		0.61
10	Most of my friends and people I know on social networks are happier than I am	0.72		0.57	
11	I compare myself with other people who I think are better than me on social networks	0.89		0.73	
12	I worry that my posts will not have enough positive interactions	0.76		0.64	
13	I worry a lot about what people might think of my content	0.77		0.68	
14	I feel lonely when I see what my contacts post on social networks	0.80		0.65	
15	My academic or work performance has declined because of using social networks		0.76		0.58
16	People close to me have complained because I use social networks too much		0.71		0.59
17	I feel that if I am not connected to social networks, I am missing out on something		0.74		0.60
18	Using social networks, I lose track of time and ignore important tasks I have outstanding		0.81		0.69
CFI		0.95			
TLI		0.94			
$RMSEA_{(90\% \ CI)}$		$0.07_{(0.068-0.079)}$			

Note. n = 698; *F.L.* factor loadings; *CFI* comparative fit index; *TLI* Tucker-Lewis index; *RMSR* root mean square of residuals; *RMSEA* $_{(90\% \ CI)}$ root mean square error of approximation (90% confidence intervals); DI^1 discrimination index of the items in the first factor, called negative social comparison; DI^2 discrimination index of the items in the second factor, called addictive consequences

Table 4. Pearson correlations between the negative social comparison subscale, addictive consequences subscale, the HADS questionnaire, and the SWLS scale

	Anxiety	Depression	Satisfaction
Comparison	0.46**	0.46**	−0.35**
Addictive	0.22**	0.21**	−0.15**

**$p \leq 0.001$

Note: $N = 1003$; Anxiety: the HADS subscale measuring anxiety; Depression: the HADS subscale measuring depression; Satisfaction: satisfaction with life scale; Comparison: the negative social comparison on SNs subscale from the problematic use of SNs scale (PUS); Addictive: the addictive consequences subscale from the problematic use of SNs scale (PUS)

Table 5. Norm-referenced scores for the negative social comparison subscale divided by sex

Men		Women	
Direct score	Percentile	Direct score	Percentile
10	20	10	15
11	24	11	20
12	37	12	27
13	49	13	33
14	53	14	39
15	56	15	42
16	61	16	47
17	64	17	54
18	66	18	58
19	69	19	61
20	72	20	70
21	80	21	78
22	82	22	79
23	85	23	83
24–25	87	24	85
26	93	25	86
27	94	26	88
28	95	27	89
29	96	28	92
30	97	29	93
31	98	30	94
32–40	99	31–32	95
		33–34	96
		35	98
		36–40	99

SNs are used in a comparative way; and (ii) it assesses the problematic use of SNs without focusing on any specific social network, and thus, it can be generalized to different SNs.

In terms of internal structure, the new PUS scale demonstrated a bidimensional structure with two strongly correlated dimensions. Therefore, on the one hand, it is possible

Table 6. Norm-referenced scores for the addictive consequences subscale divided by sex

Men		Women	
Direct score	Percentile	Direct score	Percentile
8	14	8	17
9	16	9	23
10	21	10	30
11	30	11	37
12	32	12	42
13	36	13	47
14	39	14	53
15	47	15	57
16	56	16	61
17	60	17	67
18	72	18	75
19	75	19	76
20	78	20	82
21	79	21	86
22	86	22	88
23	88	23	90
24	90	24	92
25	92	25	93
26	97	26	96
27–30	98	27	97
31–40	99	28	98
		29	98
		30–40	99

to obtain a score for each of the subscales for more specific information about problematic use of SNs. On the other hand, considering the high correlation of the subscales, a total problematic use score could also be obtained. One of the subscales, addictive consequences, can be considered similar to other scales that evaluate the addictive use of SNs in order to compare its structure. Most of the scales that assess this addictive use of SNs are unidimensional, such as the Social Media Disorder Scale (Van Den Eijnden et al., 2016), the Facebook Intrusion Questionnaire (FIQ) from Elphinston and Noller (2011) and the Bergen Facebook Addiction Scale (Andreassen et al., 2012) which focuses on Facebook, as well as its adaptation for SNs in general, the Bergen Social Media Addiction scale (Andreassen et al., 2016). However, there are other scales that evaluate this kind of SNs use with multidimensional factor structures, such as the TDI-RS from Chóliz et al. (2016).

The second factor (i.e. negative social comparison) is more difficult to compare, as there are no existing questionnaires that assess the same concept. This factor refers to the upward social comparison that could take place in the use of SNs and that may be detrimental to psychological well-being (Verduyn et al., 2020). Pre-existing questionnaires that have assessed social comparison on SNs are of two types: (i) the ones developed exclusively for a specific study with no intention of being validated and (ii) the ones adapted from questionnaires addressing social comparison in general which have been reformulated a fit a SNs environment. This latter group includes the questionnaire

created in the 2017 study by Liu et al. in which the authors adapted a questionnaire for assessing social comparison in a day-to-day non-virtual context so that it could be used to assess upward comparisons on SNs. In the study, the authors performed a factorial analysis that showed a unidimensional structure, as did Hanna et al. (2017).

Our new PUS instrument demonstrated excellent reliability according to the European model of test quality assessment (Muñiz, 2018), both in the negative social comparison ($\alpha = 0.94$; $\omega = 0.94$) and addictive consequences ($\alpha = 0.91$; $\omega = 0.91$) subscales. The latter demonstrated greater reliability than other questionnaires that measure the addictive use of SNs, such as the short version of the Social Media Disorder Scale (Van Den Eijnden et al., 2016), the Facebook Intrusion Questionnaire (Elphinston & Noller, 2011), the Bergen Facebook Addiction Scale (Andreassen et al., 2012) and the Bergen Social Media Addiction Scale (Andreassen et al., 2016). In addition, none of the items in the PUS scale exhibited DIF based on sex, and all items had high discriminatory power. It is not possible to compare the reliability of the negative social comparison subscale with validated questionnaires that measure negative comparative use of SNs as we did not find any. However, it exhibited notably higher reliability than the "ad hoc" questionnaires used by other authors (Cramer et al., 2016; Schmuck et al., 2019; Tandoc et al., 2015).

In terms of evidence of validity related to other variables, the two PUS subscales demonstrated strong positive correlations with both anxiety and depression and a negative correlation with life satisfaction. Furthermore, the canonical correlation also indicates the relationship between these two subsets. This is consistent with previous studies that have found a relationship between the use of SNs for comparisons and anxiety and depression (Hussain & Griffiths, 2018) as well as with dissatisfaction with life (Verduyn et al., 2017). Also, it is consistent with previous studies that have found a relationship between the addictive use of SNs and anxiety and depression (Barbar et al., 2021; Keles et al., 2020; Malaeb et al., n.d.; Seabrook et al., 2016; Youssef et al., 2021) and lower levels of psychological well-being (Arias-de la Torre et al., 2020; Huang, 2017).

It is worth highlighting the differences between the PUS subscales. The negative social comparison subscale correlated more strongly with depression, anxiety and dissatisfaction with life than the addictive consequences subscale. Consequently, this may indicate the importance of evaluating the social comparison carried out in SNs in the evaluation of the problematic use of SNs in addition to its addictive consequences. This importance of social comparison, even over and above that of addictive traits, has already been highlighted by other studies (Keles et al., 2020).

The present study comes with some limitations. Firstly, the sampling was not strictly random, which means that any generalization will be limited. In addition, the sample is mainly comprised of highly qualified women. In order to address the latter, weighting to population has been calculated, but it is still a limitation to be taken into consideration. We did not obtain data from under 18 years old, which means that it might be useful to replicate this study with minors. Additionally, data was collected during the Covid-19 lockdown which may have increased problematic and non-problematic use of SNs (Brailovskaia & Margraf, 2021).

In conclusion, the present study has developed a new scale for the assessment of the problematic use of SNs. The instrument is composed of 18 items divided into two subscales: addictive consequences and negative social comparison. The instrument demonstrated excellent psychometric properties. This new tool will allow a more complete assessment of the inappropriate use of SNs by not only assessing the consequences of such use but also the tendency to use of SNs to make interpersonal comparisons. In addition, we

have created two normative percentile scales differentiating between sexes in each of the two subscales.

Declarations

The study was not explicitly reviewed by an Ethics Committee, given that this is not required by our University of Oviedo, nor by the national guidelines established in the Code of Ethics of the Spanish Psychological Association. There are several reasons why an explicit approval by an Ethics Committee was not necessary: the participants evaluated were adults, the evaluation was voluntarily accepted, that is, an implicit informed consent is assumed, and the data is treated anonymously and confidentially. In addition, all the recommendations established in ISO-10667 Standard for the evaluation of people were strictly followed. The whole evaluation process and the use of the measuring instruments were carried out always following the Deontological Code of the Spanish Psychological Association (2010), as well as the International Test Commission Guidelines for Test Use (2013).

Informed Consent Informed consent was obtained from all patients for being included in the study.

Conflict of Interest The authors declare no competing interests.

References

Allahverdi, F. Z. (2021). The relationship between the items of the social media disorder scale and perceived social media addiction. *Current Psychology*, 1-8. https://doi.org/10.1007/s12144-020-01314-x

American Educational Research Association [AERA], American Psychological Association [APA] & National Council on Measurement in Education [NCME]. (2014). *Standards for the development and use of tests: The Standards for Educational and Psychological Testing*. American Educational Research Association.

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., 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

Appel, H., Crusius, J., & Gerlach, A. L. (2015). Social comparison, envy, and depression on Facebook: A study looking at the effects of high comparison standards on depressed individuals. *Journal of Social and Clinical Psychology, 34*(4), 277–289. https://doi.org/10.1016/j.copsyc.2015.10.006

Arias-de la Torre, J., Puigdomenech, E., García, X., Valderas, J. M., Eiroa-Orosa, F. J., Fernández-Villa, T., Molina, A. J., Martín, V., Serrno-Blanco, A., Alonso, J., & Espallargues, M. (2020). Relationship between depression and the use of mobile technologies and social media among adolescents: Umbrella review. *Journal of Medical Internet Research, 22*(8), e16388. https://doi.org/10.2196/16388

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. *PLoS One, 12*(1), e0169839. https://doi.org/10.1371/journal.pone.0169839

Barbar, S., Haddad, C., Sacre, H., Dagher, D., Akel, M., Kheir, N., Salameh, P., Hallit, S., & Obeid, S. (2021). Factors associated with problematic social media use among a sample of Lebanese adults: The mediating role of emotional intelligence. *Perspectives in Psychiatric Care, 57*(3), 1313–1322. https://doi.org/10.1111/ppc.12692

Blanca, M. J., & Bendayan, R. (2018). Spanish version of the Phubbing Scale: Internet addiction. Facebook intrusion, and fear of missing out as correlates. *Psicothema, 30*(4), 449–454. https://doi.org/10.7334/psicothema2018.153

Brailovskaia, J., & Margraf, J. (2021). The relationship between burden caused by Coronavirus (Covid-19), addictive social media use, sense of control and anxiety. *Computers in Human Behavior, 119*, 106720. https://doi.org/10.1016/j.chb.2021.106720

Calderón, G. C., Navarro, G. D., Lorenzo, S. U., & Ferrando, P. P. (2019). Multidimensional or essentially unidimensional? A multi-faceted factor-analytic approach for assessing the dimensionality of tests and items. *Psicothema, 31*(4), 450. https://doi.org/10.7334/psicothema2019.153

Caplan, S. E. (2010). Theory and measurement of generalized problematic Internet use: A two-step approach. *Computers in Human Behavior, 26*(5), 1089–1097. https://doi.org/10.1016/j.chb.2010.03.012

Charter, R. A. (2003). A breakdown of reliability coefficients by test type and reliability method, and the clinical implications of low reliability. *The Journal of General Psychology, 130*(3), 290–304. https://doi.org/10.1080/00221300309601160

Chóliz, M., Marco, C., & Chóliz, C. (2016). *AdicTecno: Programa de prevención de las adicciónes tecnológicas*. TEA Ediciones.

Cramer, E. M., Song, H., & Drent, A. M. (2016). Social comparison on Facebook: Motivation, affective consequences, self-esteem, and Facebook fatigue. *Computers in Human Behavior, 64*, 739–746. https://doi.org/10.1016/j.chb.2016.07.049

Cronbach, L.J (1951) Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/BF02310555

De Lenne, O., Vandenbosch, L., Eggermont, S., Karsay, K., & Trekels, J. (2018). Picture-perfect lives on social media: A cross-national study on the role of media ideals in adolescent well-being. *Media Psychology, 23*(1), 52–78. https://doi.org/10.1080/15213269.2018.1554494

De Vries, D. A., & Kühne, R. (2015). Facebook and self-perception: Individual susceptibility to negative social comparison on Facebook. *Personality and Individual Differences, 86*, 217–221. https://doi.org/10.1016/j.paid.2015.05.029

De Vries, D. A., Möller, A. M., Wieringa, M. S., Eigenraam, A. W., & Hamelink, K. (2018). Social comparison as the thief of joy: Emotional consequences of viewing strangers' Instagram posts. *Media Psychology, 21*(2), 222–245. https://doi.org/10.1080/15213269.2016.1267647

Deontological Code of the Spanish Psychological Association (2010). Madrid: Consejo General de Colegios Oficiales de Psicólogos (www.cop.es). Accessed 15 Mar 2021

Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71–75.

Downing, S. M., & Haladyna, T. M. (2006). *Handbook of test development*. Lawrence Erlbaum Associates.

Elphinston, R. A., & Noller, P. (2011). Time to face it! Facebook intrusion and the implications for romantic jealousy and relationship satisfaction. *Cyberpsychology, Behavior and Social Networking, 14*(11), 631–635. https://doi.org/10.1089/cyber.2010.0318

Evers, A., Muñiz, J., Hagemeister, C., Høstmælingen, A., Lindley, P., Sjöberg, A., & Bartram, D. (2013). Assessing the quality of tests: Revision of the EFPA review model. *Psicothema, 25*(3), 283–291. https://doi.org/10.7334/psicothema2013.97

Feijóo, S., Foody, M., Norman, J. O. H., Pichel, R., & Rial, A. (2021). Cyberbullies, the cyberbullied, and problematic internet use: Some reasonable similarities. *Psicothema, 33*(2), 198–205. https://doi.org/10.7334/psicothema2020.209

Ferrando Piera, P. J. (2021). Seven decades of factor analysis: From Yela to the present day. *Psicothema, 33*(3), 378–385. https://doi.org/10.7334/psicothema2021.24

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39–50. https://doi.org/10.2307/3151312

Frost, R. L., & Rickwood, D. J. (2017). A systematic review of the mental health outcomes associated with Facebook use. *Computers in Human Behavior, 76*, 576–600. https://doi.org/10.1016/j.chb.2017.08.001

Gibbons, F. X., & Buunk, B. P. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology, 76*(1), 129–142. https://doi.org/10.1037/0022-3514.76.1.129

Gilmour, J., Machin, T., Brownlow, C., & Jeffries, C. (2020). Facebook-based social support and health: A systematic review. *Psychology of Popular Media, 9*(3), 328–337. https://doi.org/10.1037/ppm0000246

Gómez-Benito, J., Hidalgo, M. D., & Zumbo, B. D. (2013). Effectiveness of combining statistical tests and effect sizes when using logistic discriminant function regression to detect differential item functioning for polytomous items. *Educational and Psychological Measurement, 73*(5), 875–897. https://doi.org/10.1177/0013164413492419

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate Data Analysis* (7th ed.). Prentice Hall.

Hanna, E., Ward, L. M., Seabrook, R. C., Jerald, M., Reed, L., Giaccardi, S., & Lippman, J. R. (2017). Contributions of social comparison and self-objectification in mediating associations between Facebook use and emergent adults' psychological well-being. *Cyberpsychology, Behavior and Social Networking, 20*(3), 172–179. https://doi.org/10.1089/cyber.2016.0247

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

Huang, C. (2017). Time spent on social network sites and psychological well-being: A meta-analysis. *Cyberpsychology, Behavior and Social Networking, 20*(6), 346–354. https://doi.org/10.1089/cyber.2016.0758

Hussain, Z., & Griffiths, M. D. (2018). Problematic social networking site use and comorbid psychiatric disorders: A systematic review of recent large-scale studies. *Frontiers in Psychiatry, 9*, 686. https://doi.org/10.3389/fpsyt.2018.00686

Hussain, Z., & Griffiths, M. D. (2021). The associations between problematic social networking site use and sleep quality, attention-deficit hyperactivity disorder, depression, anxiety and stress. *International Journal of Mental Health and Addiction, 19*, 686–700. https://doi.org/10.1007/s11469-019-00175-1

IBM Corp. (2016). *IBM SPSS Statistics for Windows*, Version 24.0. Armonk, NY: IBM Corp. [Computer software].

International Test Commission Guidelines for Test Use (2013). International Test Commission: www.intestcom.org. Accessed 15 Mar 2021

Keles, B., McCrae, N., & Grealish, A. (2020). A systematic review: The influence of social media on depression, anxiety and psychological distress in adolescents. *International Journal of Adolescence and Youth, 25*(1), 79–93. https://doi.org/10.1080/02673843.2019.1590851

Lane, S., Raymond, M. R., Haladyna, T. M., & Downing, S. M. (2015). Test development process. In S. Lane, M. R. Raymond, & T. M. Haladyna (Eds.), *Handbook of test development* (2nd ed., pp. 3–18). Lawrence Erlbaum Associates.

Lee, S. Y. (2014). How do people compare themselves with others on social network sites?: The case of Facebook. *Computers in Human Behavior, 32*, 253–260. https://doi.org/10.1016/j.chb.2013.12.009

Liu, Q. Q., Zhou, Z. K., Yang, X. J., Niu, G. F., Tian, Y., & Fan, C. Y. (2017). Upward social comparison on social network sites and depressive symptoms: A moderated mediation model of self-esteem and optimism. *Personality and Individual Differences, 113*, 223–228. https://doi.org/10.1016/j.paid.2017.03.037

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada. *Anales de Psicología, 30*(3), 1151–1169. https://doi.org/10.6018/analesps.30.3.199361

Lorenzo-Seva, U., & Ferrando, P. J. (2013). FACTOR 9.2: A comprehensive program for fitting exploratory and semiconfirmatory factor analysis and IRT models. *Applied Psychological Measurement, 37*(6), 497–498. https://doi.org/10.1177/0146621613487794

Luong, K. T., Knobloch-Westerwick, S., & Frampton, J. (2019). Temporal self impacts on media exposure & effects: A test of the Selective Exposure Self-and Affect-Management (SESAM) model. *Media Psychology, 24*(3), 48–78. https://doi.org/10.1080/15213269.2019.1657898

Malaeb, D., Salameh, P., Barbar, S., Awad, E., Haddad, C., Hallit, R., Sacre, H., Akel, M., Obeid, S., & Hallit, S. (2021). Problematic social media use and mental health (depression, anxiety, and insomnia) among Lebanese adults: Any mediating effect of stress? *Perspectives in Psychiatric Care, 57*(2), 539–549. https://doi.org/10.1111/ppc.12576

Marino, C., Gini, G., Vieno, A., & Spada, M. M. (2018a). A comprehensive meta-analysis on problematic Facebook use. *Computers in Human Behavior, 83*, 262–277. https://doi.org/10.1016/j.chb.2018.02.009

Marino, C., Gini, G., Vieno, A., & Spada, M. M. (2018b). The associations between problematic Facebook use, psychological distress and well-being among adolescents and young adults: A systematic review and meta-analysis. *Journal of Affective Disorders, 226*, 274–281. https://doi.org/10.1016/j.jad.2017.10.007

Marino, C., Vieno, A., Altoè, G., & Spada, M. M. (2016). Factorial validity of the Problematic Facebook Use Scale for adolescents and young adults. *Journal of Behavioral Addictions, 6*(1), 5–10. https://doi.org/10.1556/2006.6.2017.004

Matute, H. (2016). Adicción, abuso o uso problemático de Internet. In E. En Echeburúa (Ed.), *Abuso de Internet ¿Antesala para la adicción al juego de azar online?* (pp. 31–46). Pirámide.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Erlbaum.

Muñiz, J. (2018). *Introducción a la psicometría*. Pirámide.

Muñiz, J., Fidalgo, A. M., García-Cueto, E., Martínez, R., & Moreno, R. (2005). *Item Analysis*. La Muralla.

Muñiz, J., & Fonseca-Pedrero, E. (2019). Diez pasos para la construcción de un test. *Psicothema, 31*(1), 7–16. https://doi.org/10.7334/psicothema2018.291

Muthen, B., & Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model. *British Journal of Mathematical and Statistical Psychology, 45*(1), 19–30. https://doi.org/10.1111/j.2044-8317.1992.tb00975.x

Muthén, L. K., & Muthen, B. (2017). *Mplus user's guide: Statistical analysis with latent variables, user's guide*. Muthén & Muthén.

Penfield, R. D., & Giacobbi Jr., P. R. (2004). Applying a score confidence interval to Aiken's item content-relevance index. *Measurement in Physical Education and Exercise Science, 8*(4), 213–225. https://doi.org/10.1207/s15327841mpee0804_3

Perloff, R. M. (2014). Social Media Effects on Young Women´s Body Image Concerns: Theoretical Perspectives and an Agenda for Research. *Sex Roles, 71*(11-12), 363–377. https://doi.org/10.1007/s11199-014-0384-6

Quintana, J. M., Padierna, A., Esteban, C., Arostegui, I., Bilbao, A., & Ruiz, I. (2003). Evaluation of the psychometric characteristics of the Spanish version of the Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 107*(3), 216–221. https://doi.org/10.1034/j.1600-0447.2003.00062.x

Schmuck, D., Karsay, K., Matthes, J., & Stevic, A. (2019). "Looking Up and Feeling Down". The influence of mobile social networking site use on upward social comparison, self-esteem, and well-being of adult smartphone users. *Telematics and Informatics, 42*, 101240. https://doi.org/10.1016/j.tele.2019.101240

Seabrook, E. M., Kern, M. L., & Rickard, N. S. (2016). Social networking sites, depression, and anxiety: A systematic review. *Journal of Medical Internet Research, 3*(4), e50. https://doi.org/10.2196/mental.5842

Sireci, S., & Faulkner-Bond, M. (2014). Evidencia de validez basada en el contenido del test. *Psicothema, 26*(1), 100–108. https://doi.org/10.7334/psicothema2013.256

Suárez-Alvarez, J., Pedrosa, I., Lozano Fernández, L. M., García-Cueto, E., Cuesta, M., & Muñiz, J. (2018). Using reversed items in Likert scales: A questionable practice. *Psicothema, 30*(2), 149–158. https://doi.org/10.7334/psicothema2018.33

Tandoc, E. C., Ferrucci, P., & Duffy, M. (2015). Facebook use, envy, and depression among college students: Is facebooking depressing? *Computers in Human Behavior, 43*, 139–146. https://doi.org/10.1016/j.chb.2014.10.053

Terol, M. C., López-Roig, S., Rodríguez-Marín, J., Martí-Aragón, M., Pastor, M. A., & Reig, M. T. (2007). Propiedades psicométricas de la escala hospitalaria de ansiedad y depresión (had) en población española. *Ansiedad y Estrés, 13*(2-3), 163–176.

Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods, 16*(2), 209–220. https://doi.org/10.1037/a0023353

Vahedi, Z. & Zannella, L. (2021). The association between self-reported depressive symptoms and the use of social networking sites (SNS): A meta-analysis. *Current Psychology*, 1-16. https://doi.org/10.1007/s12144-019-0150-6

Van Den Eijnden, R. J., 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

Vázquez, C., Duque, A., & Hervás, G. (2013). Satisfaction with life scale in a representative sample of Spanish adults: Validation and normative data. *The Spanish Journal of Psychology, 16*, 1–15. https://doi.org/10.1017/sjp.2013.82

Verduyn, P., Gugushvili, N., Massar, K., Täht, K., & Kross, E. (2020). Social comparison on social networking sites. *Current Opinion in Psychology, 36*, 32–37. https://doi.org/10.1016/j.copsyc.2020.04.002

Verduyn, P., Ybarra, O., Résibois, M., Jonides, J., & Kross, E. (2017). Do social network sites enhance or undermine subjective well-being? A critical review. *Social Issues and Policy Review, 11*(1), 274–302. https://doi.org/10.1111/sipr.12033

Vigil-Colet, A., Navarro-González, D., & Morales-Vives, F. (2020). To reverse or to not reverse Likert-type items: That is the question. *Psicothema, 32*(1), 108–114. https://doi.org/10.7334/psicothema2019.286

Yoon, S., Kleinman, M., Mertz, J., & Brannick, M. (2019). Is social network site usage related to depression? A meta-analysis of Facebook-depression relations. *Journal of Affective Disorders, 248*, 65–72. https://doi.org/10.1016/j.jad.2019.01.026

Youssef, L., Hallit, R., Akel, M., Kheir, N., Obeid, S., & Hallit, S. (2021). Social media use disorder and alexithymia: any association between the two? Results of a cross-sectional study among Lebanese adults. *Perspectives in Psychiatric Care, 57*(1), 20–26. https://doi.org/10.1111/ppc.12506

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67*(6), 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.