

ASSESSING ADOLESCENTS' INTERNALIZING SYMPTOMS USING VIRTUAL SOCIAL NETWORKS: A FORMAT EQUIVALENCE STUDY OF THE REVISED CHILD ANXIETY AND DEPRESSION SCALE

Alejandro de la Torre-Luque^{1,2}, Maria Balle³, Aina Fiol-Veny³,
Xavier Bornas³ and Albert Sese³

¹Universidad Autónoma de Madrid; ²Centro de Investigación Biomédica en Red en Salud Mental (CIBERSAM); ³University of the Balearic Islands (Spain)

Abstract

The widespread use of virtual social networks (VSN) by adolescents makes it possible to conduct psychological assessments or health promotion using these platforms. However, psychometric properties of them should be validated. This study aimed to test the feasibility of administering the Revised Child Anxiety and Depression Scale (RCADS) over a secure social network, while preserving its original psychometric properties. To do so, a sample of 703 adolescents ($M= 13.86$ years, $SD= 0.49$) completed the questionnaire either over a social network or using paper and pencil. We tested a two-way format equivalence: quantitative equivalence (comparing score distributions across versions); and qualitative or conceptual equivalence (comparing the between-factor correlations between versions and measurement invariance). As a result, no difference was found between the score distributions of the two versions and between-factor intercorrelations with similar patterns in both versions. Finally, both methods for administering the RCADS showed an adequate fit with their theoretical latent structure, thus preserving format equivalence. To sum up, VSN may, therefore, constitute appropriate contexts for conducting psychological assessment and research among adolescents.

KEY WORDS: *virtual social networks, format equivalence, psychological assessment, adolescence.*

Resumen

El uso de las redes sociales por parte de adolescentes está bastante extendido actualmente. Esto posibilita el desarrollo de programas de evaluación y promoción de la salud mediante dichas vías, si se conservan garantías psicométricas en su uso. Este estudio pretendía comprobar si la "Escala revisada de ansiedad y depresión infantil" (RCADS) mantenía sus propiedades psicométricas originales cuando era aplicado mediante una red social. Una muestra de 703 adolescentes ($M= 13,86$ años; $DT= 0,49$) completaron la escala en formato tradicional o en una red social. Se evaluaron dos tipos de equivalencia entre formatos: equivalencia cuantitativa (distribución de las puntuaciones entre formatos) y cualitativa o conceptual

(patrones de intercorrelaciones e invarianza de medida entre formatos). No se encontraron diferencias de las escalas en el formato tradicional y en una red social. Además, el patrón de intercorrelaciones entre factores fue similar y se observó invarianza de medida entre formatos. En conclusión, la versión en red social de la RCADS mostró propiedades psicométricas equivalentes a la tradicional, destacándose la aplicabilidad en estas plataformas.

PALABRAS CLAVE: *redes sociales, equivalencia de formatos, evaluación psicológica, adolescencia.*

Introduction

Nowadays, virtual social networking has become fashionable, and it is available to everyone. Furthermore, these networks are good sources for gathering research samples as they offer access to large communities of people and allow patient follow-ups to be managed more easily (Ilioudi, Lazakidou, Glezakos, & Tsironi, 2012; Park & Calamaro, 2013). On this basis, incorporating computerized tools for assessing internalizing symptoms and health-related issues into the functionalities of social networking platforms seems like it would be very useful. This may allow for monitoring symptomatology or preventing the development of full-blown disorders from at-risk conditions. In this vein, factors highly correlated with the development of internalizing disorders, such as some temperamental traits (e.g., high negative affectivity or low effortful control; see García-López, Piqueras, Díaz-Castela, & Inglés, 2008; Rothbart, 2007; Rothbart & Gartstein, 2008) have already been assessed using computer-based instruments (see Putnam & Stifter, 2008). Computerized versions of psychometric instruments have also been utilized in the assessment of psychopathology among adolescents (Fonseca-Pedrero, Fernández, Paino, Lemos-Giraldez, & Muñiz, 2013; Hamann, Schultze-Lutter, & Tarokh, 2016; Van Ballegooijen Riper, Cuijpers, Van Oppen, & Smit, 2016). Several studies have affirmed that traditionally administered instruments and those using computer devices or the Internet are equivalent in terms of psychometric properties for adolescent populations (Richardson, Johnson, Ratner, & Zumbo, 2009; Wyrick & Bond, 2011). However, some other studies have pointed to an absence of format equivalence between these different methods of administering the same questionnaire (Hirai, Vernon, Clum, & Skidmore, 2011; Patalay, Hayes, Deighton, & Wolpert, 2016).

Ensuring the equivalence of formats is a vital task that must be undertaken before employing computer-based assessment in order to avoid biased and misrepresented conclusions. Validity-related discrepancies between versions may emerge due to the study design, the data collection processes, or how the assessments are administered (Green, 1991; Weigold, Weigold, & Russell, 2013).

In this study, the format equivalence between two versions of the same internalizing symptom questionnaire, used widely for adolescents, was investigated (Donnelly, Fitzgerald, Shevlin, & Dooley, 2018; Piqueras, Martín-Vivar, Sandín, San Luis, & Pineda, 2017; Piqueras, Pineda, Martín-Vivar, & Sandín, 2017; Stevanovic et al., 2017). One was a pencil and paper version, and the other was a version that was administered through a social network. To our best knowledge, the

psychometric properties of a computerized version from one of the widely used questionnaire to assess internalizing symptoms in adolescence, the Revised Child Anxiety and Depression Scale (RCADS; Chorpita, Yim, Moffitt, Umemoto, & Francis, 2000), have not been studied yet.

We followed a two-way approach: first, studying the equivalence of the score distributions of both versions and, second, analyzing whether the latent constructs and their structural relationship to items remained invariant across versions. The general hypothesis was that both formats would be equivalent, i.e., the same constructs and their representations would be maintained regardless of which version was administered. Specifically, we expected that: (a) differences between the score distributions from both versions (formats) of the questionnaire would not be statistically significant, (b) the same pattern of factor intercorrelations would be found in both versions, (c1) both the pencil and paper and the social-network versions would show factor structures in line with the questionnaire's theoretical factor structure, and (c2) the factor structure of both versions of each of the two administered questionnaires would be equivalent.

Method

Participants

Our initial sample comprised 932 high-school students ($M= 13.92$ years; $SD= 0.56$; 45.40% males) from grades 1 and 2, who were fluent in both Spanish and Catalan and had consented to participating in writing, signed by themselves and their legal guards. According to parent and school board reports, none of them was diagnosed with any intellectual disability disorder or any neurological or developmental disorder (American Psychiatric Association, 2000).

Instruments

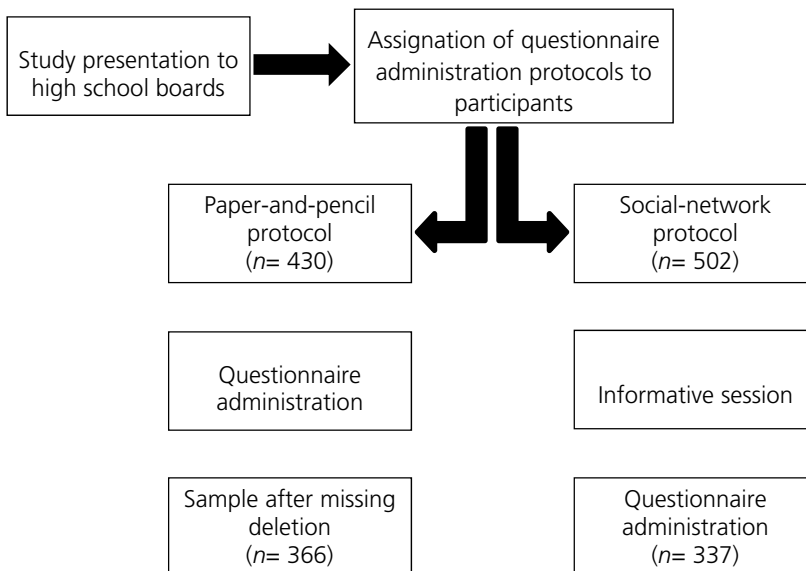
- a) *Ad hoc Socio-demographic Questionnaire*. This instrument presented respondents with questions related to their age and family composition.
- b) *Early Adolescence Temperament Questionnaire Revised* (EATQ-R; Ellis & Rothbart, 2001), Catalan version by Mary Rothbart's Temperament Lab (2015). Instrument consisting of 103 self-report items to explore temperamental features. Items in this questionnaire are scored on a 5-point Likert scale. The EATQ-R shows the factor structure of 13 dimensions of temperament as well as four second-order constructs (negative affectivity, effortful control, affiliation, and surgency). This questionnaire was administered to control for temperamental factors between study groups. We found similar levels of reliability between reference studies (see Ellis & Rothbart, 2001; Muris & Meesters, 2009) and EATQ-R scales in our sample (Cronbach's α from .50 to .80 across scales, for the social-network version; and Cronbach's α from .50 to .83 across scales, for the pencil and paper version).
- c) *Revised Child Anxiety and Depression Scale* (RCADS; Chorpita et al., 2000), Spanish version by Sandín, Valiente, and Chorot (2009). Instrument composed

of 47 items with a 4 point Likert-type scale of response to assess symptomatology of anxiety disorders and major depression in childhood and adolescence. The RCADS has a correlated 6-factor structure (panic disorder, social phobia, separation anxiety, obsessive-compulsive disorder, generalized anxiety disorder, and major depression). In this study, we used the Spanish version of the RCADS, as no version in regional language exists. Reliability levels were appropriate for the Spanish version, with Cronbach's alpha between .68 to .81, for the scales (Sandin et al., 2009). Chorpita, Moffitt, and Gray (2005) found the cutoff points for clinical meaningfulness, being for the separation anxiety scale= 5; social phobia= 10; generalized anxiety= 7; panic disorder symptomatology= 12; obsessive compulsive disorder= 5; and major depression= 11. These scores should be considered in caution as they were based on U.S. samples.

Procedure

Initially, the study was presented to 20 secondary school boards (all of the state schools, 45% from rural areas) in Majorca (Spain). Afterwards, every school was assigned to one of the study protocols (Figure 1), 13 of them voluntarily accepted to undergo the social-network protocol and seven the pencil and paper protocol. One information session was held at each school to provide the students with information and encourage individual participation.

Figure 1
Study protocols of sample recruitment and questionnaire administration



THE PENCIL AND PAPER PROTOCOL

Researchers attended each secondary school and gathered all participants in a classroom. A brief explanation of how questionnaires should be filled in was provided. The questionnaires (RCADS and EATQ-R) were administered in two non-consecutive 1-hour sessions in counterbalanced order. Prior to responding to the questionnaires, the sociodemographic self-report was given.

Once the questionnaires were collected, data derived from them were electronically processed by optical character recognition devices at a university. Questionnaires missing more than 5% of values were excluded. Missing values for the non-excluded questionnaires were estimated by means of a multiple imputation procedure with a threshold of five iterations (see Brand, 1999). Finally, item scores were calculated along with the psychometric properties elucidated by the questionnaires.

THE SOCIAL-NETWORK PROTOCOL

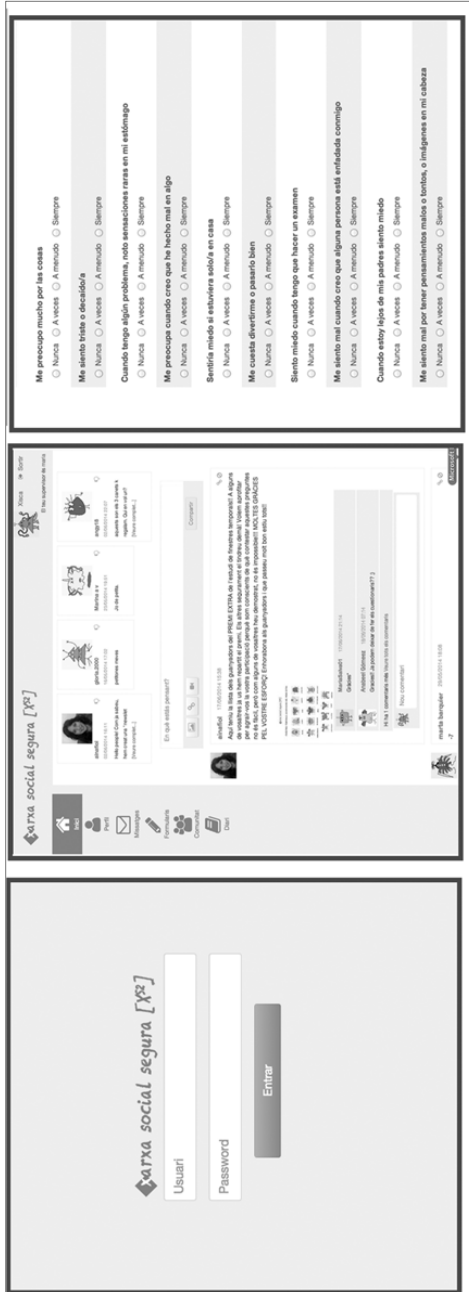
XS2 (Xarxa Social Segura [Secure Social Network] -www.xs2.cat) is a secure social network developed for research purposes within a wider research project (Bornas, Llabres, Balle, de la Torre-Luque, & Fiol-Veny, 2014). It provides an environment for secure interaction among adolescents, secondary school teachers, and researchers. Among its other features, it allows researchers to conduct psychological follow-ups and to apply psychometric instruments that adolescents can fill in from anywhere with an Internet connection. Furthermore, teachers can get involved in social networking with the adolescents, thus encouraging the appropriate use of these types of platforms (e.g., promoting avoiding offensive and sexist language and the posting of inappropriate content). Moreover, XS2 allows adolescents to satisfy their social needs for relationships and expression through virtual social networks. Some XS2 features are shown in Figure 2.

Participants assigned to the social-network protocol had already used XS2 satisfactorily before the questionnaires were activated. An alert symbol on the participants' XS2 wall page informed them that the questionnaires were available to be filled in until they were completed. Each participant could freely choose which one to complete first, when to do it and where (just if they had access to the internet and the social network site). This protocol required participants to answer all of the items on each questionnaire (by clicking the option more appropriate to them), i.e., no missing values were permitted. A one-month deadline was set for completing the questionnaires.

Data analysis

Two types of format equivalence were analyzed (Green, 1991; Weigold et al., 2013). Concerning our first hypothesis (a), we evaluated the quantitative equivalence by testing if the score distributions of the RCADS factors were similar in both versions. These distributions between versions were compared by means of *t*-tests for independent samples. In addition, the 20% equivalence interval was

Figure 2
Different functionalities of the secure social network XS2



Note. The XS2 initial logging page is presented at left side; the “wall” or principal page of the social network at the middle; and a page where each participant responded to the RCADS at right side.

calculated (Steele, Mummery, & Dwyer, 2009). This determines the boundaries by which a difference between the means of two groups is meaningful. Thus, a difference falling outside of the 20% interval from the mean of the reference group (in our case, the pencil and paper protocol group) would reflect a lack of meaningful quantitative equivalence between versions.

Conceptual equivalence was studied by testing the other study hypotheses. Regarding our second hypothesis (b), intercorrelations between questionnaire factors were compared across versions by using z-tests based on Fisher's transformations (Cohen & Cohen, 1983).

Testing for the c1 and c2 hypotheses involved examining construct validity. We assumed that if construct validity were preserved, constructs of interest could be measured in the same way across groups (versions). We therefore aimed to test for measurement invariance (MI) in the responses to both versions of the RCADS (see Meredith & Teresi, 2006) via a multi-group confirmatory factor analysis (MGCFA; Billiet, 2002; Jöreskog, 1971). Thus, the fit of the single-group solutions (the pencil and paper version and the social-network one, independently) to their theoretical models was tested first.

Afterwards, the MGCFA pooled the data from both RCADS versions to test for MI. This was done by following a stepwise strategy. The process started by examining the unconstrained solution of the pooled data (the so-called configural invariance) followed by increasingly restrictive models: constrained factorial weights (metric or weak MI), an additional constraint on item intercepts (scalar or strong MI), and finally, constraints were added to error variances and covariances (strict MI).

The MGCFA parameters were estimated using weighted least squares means and variance adjusted (WLSMV) methods because multivariate normality could not be assumed for the RCADS (Beauducel & Herzberg, 2006).

Five indices were selected to test model fit (Thompson & Green, 2006): the χ^2 statistic, for goodness-of-fit; the root mean square error of approximation (RMSEA) index, with a 90% confidence interval; the standardized root mean square residual (SRMR); the comparative fit index (CFI); the goodness-of-fit index (GFI); and the Tucker-Lewis index (TLI). Good fit is proven when $RMSEA \leq .05$ (at a 90% confidence interval with a lower bound no larger than .05 and an upper bound smaller than .08), $SRMR < .08$, $CFI \geq .95$, $GFI \geq .95$, and $TLI \geq .95$. Additionally, comparisons between MI-based nested models were conducted by means of χ^2 -based test (Satorra & Bentler, 2001) and incremental CFI (ΔCFI). These statistics were calculated between the two proximal models (unconstrained vs. weak MI model; weak MI vs. strong MI model, etc.). If a χ^2 -based test reflected significant differences between these proximal models, the simpler one was not nested within the more complex one, and therefore, MI could not be upheld. Moreover, if $\Delta CFI > -.002$ between two proximal models, some evidence of a lack of MI would be provided (Meade, Johnson, & Braddy, 2008).

All analyses were conducted using IBM SPSS Statistics 20.0 (IBM, 2011) and R x64 3.0.1 (LAVAAN package) (Rosseel, 2012).

Results

A total of 430 adolescents were assigned to the pencil and paper protocol. Out of these, 366 participants filled out the RCADS and had less than 5% of values missing (rate of participation= 85.12%). Additionally, 43 out of 17202 values were imputed for the questionnaires. A total of 502 adolescents were assigned to the social-network protocol, and 337 of them (rate of participation= 67.13%) filled out the RCADS (Table 1). As a result, data from 703 adolescents ($M= 13.86$ years; $SD= .49$) were analyzed.

Reliability indexes for the RCADS were appropriate (Cronbach's α from .73 to .87 across scales, for the social-network version; and Cronbach's α from .74 to .84 across scales, for the pencil and paper version).

Table 1
Sample distribution and demographic features

Variables	Paper-and-pencil version	Social-network version
<i>N</i>	366	337
Gender (males)	42.30	42.40
Age**	13.96 (.58)	13.76 (.40)
Family structure		
Parents and siblings	62.40	77.40
Parents and no siblings	11.60	3.20
Divorced parents and siblings	13.60	11.30
Divorced parents and no siblings	11.30	6.50
Others family structures	1.20	1.60
Adolescence Temperament Questionnaire-Revised (EATQ-R)		
Affiliativeness**	3.54 (.58)	3.42 (.52)
Effortful control	4.42 (.49)	4.41 (.54)
Negative affectivity	3.00 (.60)	3.04 (.54)
Surgency	4.39 (.51)	4.45 (.47)

Notes: Categorical variables are displayed as percentage of cases. Age and EATQ-R scores are represented as means and standard deviations (between brackets). * $p < .05$; ** $p < .01$.

Quantitative equivalence

Significant between-group differences were shown for age (older adolescents assigned to the paper-and-pencil condition) and the temperamental factor affiliativeness (higher levels in the paper-and-pencil condition participants) (Table 1). Association analyses were conducted to account for whether or not age-related or affiliation-related differences between samples influenced scores on the questionnaire. However, these variables were shown not to have any significant influence on criteria.

Analyses derived from *t*-tests did not show significant differences between versions for any RCADS factor (Table 2). Additionally, all factor discrepancies

between the different versions fell within the equivalence interval, suggesting a lack of meaningful difference between versions.

Table 2

Score distributions of responses to both the Revised Child Anxiety and Depression Scale (RCADS) versions and their comparisons

RCADS factors	Paper-and-pencil version		Social-network version		<i>t</i>	<i>df</i>	Mean difference	95% CI	20% EI
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>					
Major depression	8.26	5.60	7.87	5.38	-.95	700.00	-.40	[-1.21, -.42]	+/- 1.65
Panic disorder	5.54	4.72	5.48	4.96	-.15	700.00	-.06	[-.77, .66]	+/- 1.11
Social phobia	9.95	5.37	10.01	5.94	.14	678.42	.06	[-.78, .89]	+/- 1.99
Separation anxiety	2.02	2.69	2.04	2.92	.11	700.00	.02	[-.39, .44]	+/- .40
Generalized anxiety disorder	6.49	3.59	6.37	3.50	-.42	701.00	-.12	[-.65, .40]	+/- 1.30
Obsessive-compulsive disorder	4.00	3.40	3.79	3.28	-.83	700.00	-.21	[-.71, .28]	+/- .80

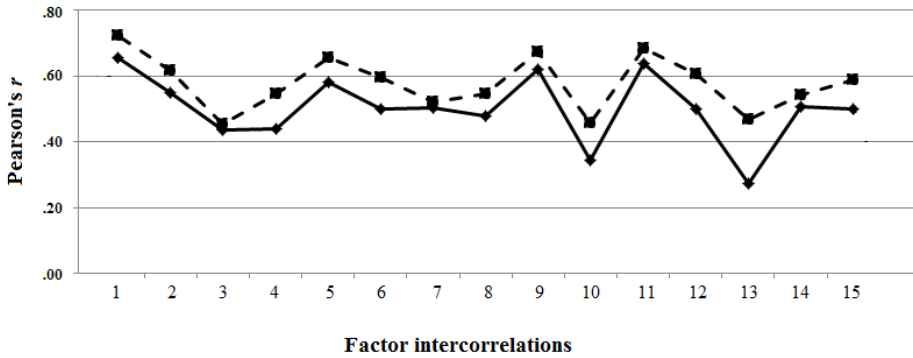
Notes: It was shown no significant comparisons between factors for both RCADS versions. EI= equivalence interval boundaries based on the reference group (paper-and-pencil version).

Conceptual equivalence

As for the intercorrelations between factors, similar patterns for both versions can be seen in Figure 3. A significantly different correlation was found for two RCADS factors (generalized anxiety symptomatology and separation anxiety symptomatology) across versions, with $z = -2.53$, $p < .01$. However, these intercorrelations were in the same direction.

Regarding the factorial solutions for each version of the RCADS, acceptable fits with their theoretical structures were shown according to χ^2 tests; RMSEA < .05, SRMR < .08, and values above .95 for CFI, GFI, and TLI (see Table 3).

Figure 3
Patterns of intercorrelation distributions across the Revised Child Anxiety and Depression Scale (RCADS) versions



Notes: Between-factor intercorrelations are expressed by using the Pearson's product-moment coefficient. Dashed line depicts the paper-and-pencil intercorrelations. Solid line depicts the social-network intercorrelations. Factor intercorrelations were: 1= Major depression vs. Panic disorder; 2= Major depression vs. Social phobia; 3= Major depression vs. Separation anxiety; 4= Major depression vs. Generalized anxiety; 5= Major depression vs. Obsessive compulsive disorder; 6= Panic disorder vs. Social phobia; 7= Panic disorder vs. Separation anxiety; 8= Panic disorder vs. Generalized anxiety; 9= Panic disorder vs. Obsessive compulsive disorder; 10= Social phobia vs. Separation anxiety; 11= Social phobia vs. Generalized anxiety; 12= Social phobia vs. Obsessive compulsive disorder; 13= Separation anxiety vs. Generalized anxiety (significantly different, with $p < .01$); 14= Separation anxiety vs. Obsessive compulsive disorder; 15= Generalized anxiety vs. Obsessive compulsive disorder.

Concerning MI, both the MGCF and model comparisons suggested equivalence between versions. First, both the unconstrained and constrained models showed good fit indices. Second, the unconstrained model was not significantly different than the proximal one (weak MI model). Finally, it was observed that the ΔCFI score was not higher than -.002 for this comparison, suggesting that a lack of measurement invariance could not be supported (Table 3).

Table 3
Model comparison summary for qualitative equivalence testing across the Revised Child Anxiety and Depression Scale (RCADS) versions

Model	χ^2 (df)	RMSEA [90% CI]	SRMR	CFI	TLI	GFI	Comparisons between MI models	
							χ^2 (df)	Δ CFI
Single solutions								
Paper-and-pencil	1144.41 (1019)	.018 [.011-.024]	.062	.99	.99	.99		
Social-network	965.08 (1019) ^{ns}	.000 [.000-.008]	.060	1.00	1.00	.98		
Multigroup solution								
Unconstrained	2109.49 (2038) ^{ns}	.010 [.000-.016]	.061	.99	.99	.97	--	--
Weak MI	2279.97 (2079)	.017 [.011-.021]	.063	.99	.99	.97	37.50 (41) ^{ns}	-.002
Strong MI	2324.05 (2120)	.017 [.011-.021]	.064	.99	.99	.97	48.77 (41) ^{ns}	.000
Strict MI	2390.51 (2167)	.017 [.011-.021]	.066	.99	.99	.97	55.65 (44) ^{ns}	-.001

Notes: MI= measurement equivalence; RMSEA= root mean square error of approximation index (scores below .05 depict good model fitting); SRMR= standardised root mean square residual (scores below .08 depict good model fitting); CFI= comparative fit index; TLI= Tucker-Lewis index; GFI= goodness-of-fit index. Scores of .95 or more indicate good model fitting, for TLI, CFI and GFI. Δ CFI= incremental CFI (scores above .002 depict poorer fit for less constrained model). All χ^2 -based tests were significant with a $p < .05$, but not in those case with superindex ^{ns} (not significant). All the comparisons between MI models were conducted between the two proximal ones. The unconstrained models keep no restrictions across versions; the weak MI models constrain the factorial loadings; the strong MI models add constraints on item intercepts; and the strict MI models include the error variances and covariances equally constrained.

Discussion

This study aimed to demonstrate that a widely used questionnaire could be administered using a secure social network while retaining the same psychometric guarantees as the corresponding traditional pencil and paper version. Hence, we conducted a format equivalence study following a two-way approach: the first one was related to quantitative score distributions and the second was conceptual and focused on construct validity.

The results of this study confirm that the format equivalence across the two versions of the RCADS was preserved. As such, relatively equivalent score distributions were observed in both versions, i.e., participants who responded to the social-network version showed scores with means and scatter distributions similar to those who filled out the traditional version. Furthermore, reliability indexes and rank distributions were similar to the regulatory ranges set by the reference version (Kosters, Chinapaw, Zwaanswijk, Van der Wal, & Koot, 2015; Sandin et al., 2009).

These results suggest that participants who could choose when to answer a questionnaire through XS2 did not necessarily score differently than those who were not free to choose when they would respond to the same questionnaire using traditional methods.

Conceptual equivalence, tested by means of between-factor correlation analyses and MGCFA, provided support for the format equivalence of the two versions. The pattern of intercorrelations between RCADS factors was proven to be equivalent across versions. Furthermore, the MGCFA revealed measurement invariance between versions. It was previously demonstrated that both methods of administering the RCADS preserved the original theoretical latent structure, independently. Afterwards, we observed that the fit between the theoretical and empirical models arising from the pooled samples was equally good. In this regard, goodness-of-fit indices (RMSEA, SRMR, CFI, GFI, and TLI) confirmed the appropriateness of the unconstrained model.

Likewise, the same latent constructs and their relations were preserved across RCADS versions in spite of slight variations in model parameters. This fact is supported by the comparisons between MI models which showed that the unconstrained model was not significantly different than the weak MI model (which was not significantly different than the strong MI model, and so on) allowing us to conclude that the unconstrained model was nested within the constrained ones. Moreover, analyses based on the Δ CFI discarded a lack of measurement equivalence with regard to the unconstrained model and the proximal constraining one, as well as for the successive comparisons (weak MI model vs. strong MI model and strong MI model vs. strict MI model).

To sum up, this study indicates that it is feasible to administer psychological assessment questionnaires pertaining to internalizing disorders to adolescents via social networking sites, disorders highly prevalent among them (Caballo et al., 2008; Copeland, Angold, Shanahan, & Costello, 2014; García-López, Salvador, & de los Reyes, 2016; Kessler et al., 2012; McDonough-Caplan, Klein, & Beauchaine, 2018; Pineda, Martín-Vivar, Sandín, & Piqueras, 2018). It is worth taking into account that these sites have some features which make them highly valuable for and fashionable among adolescents. These sites also facilitate sample recruitment, clinical assessment, follow-ups, and the delivery of health interventions, even for preventive purposes.

This study is ground-breaking in terms of incorporating clinical assessment tools into social networking sites, as other studies already showed some evidence for the feasibility of successfully adding other media-related devices within clinical protocols (Diemer, Muhlberger, Pauli, & Zwanzger, 2014; Nikou & Economides, 2017; Quero et al., 2014). Moreover, it suggests that these sites could become appropriate contexts from which to study a variety of psychological matters as the study results provide some evidence for clinically relevant constructs (anxiety and depressive symptomatology).

As a limitation, this study only explored two of the format equivalences defined by the reference authors. Another way of exploring format equivalence, called auxiliary equivalence, exists, but it is not included in this study (see Weigold et al., 2013). It utilizes missing data distributions and response times. We did not test for

auxiliary equivalence in this study because the social-network version of the RCADS did not allow for missing values. This might increase the biases derived from random responses to items as well as the drop-out rate. However, we do not consider the results of this study to show inflated or underestimated scores when comparing the social-network version to the traditional pencil and paper one. Nonetheless, future researches should include larger samples and more tests to verify auxiliary equivalence.

Also, it is important to say that participants in the social-network condition filled the questionnaires in ecological conditions. This means that they could complete the questionnaires in their naturalistic context (without any supervision by the researcher or their legal guardians nor counterbalanced questionnaire administration) with the related risks for internal validity. However, generalizability of results were higher due to the ecological conditions of application. Further research should consider more ways to control for some sources of internal validity threatening on a daily basis.

Another possible issue is that an MI model based on constraining latent means was not tested. This model could provide equivalent results to t -tests (see Thompson & Green, 2006). We decided to use t -tests in keeping with the protocols established by Weigold et al. (2013).

Finally, it is worth noting the lack of a random assignment to the study conditions (secondary schools were not randomly assigned to study conditions). However, we found no significant differences between relevant features of groups (or they did not influence our dependent variables), thus illustrating the homogeneity between study groups. Nevertheless, future studies should incorporate protocols to assign groups randomly to study conditions in order to control for this potential threat to internal validity.

It would also be very useful to implement a longitudinal approach for modeling more complex interdependent relationships where changes over time are taken into account.

Social networking platforms may represent an optimal context for developing adolescent research and prevention programs because of the features they offer. As has been corroborated by this study, clinical aspects of adolescents can be explored and followed up on with psychometric tools by using the XS2 secure social network, in order to obtain comparable results to those obtained by traditional means, optimize research resources, consolidate participant follow-ups, and expand new fields of application.

References

- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling*, 13, 186-203.
- Billiet, J. (2002). Cross-cultural equivalence with structural equation modeling. In P. P. Mohler (Ed.), *Cross-cultural survey methods* (pp. 247-264). New Jersey, NJ: John Wiley & Sons.

- Bornas, X., Llabres, J., Balle, M., de la Torre-Luque, A., & Fiol-Veny, A. (2014). The use of a secure social network for the study of anxiety in adolescence. *Revista de Psicopatología y Psicología Clínica, 19*, 171-184.
- Brand, J. P. L. (1999). *Development, implementation and evaluation of multiple imputation strategies for the statistical analysis of incomplete data sets*. Enschede: Print Partners Ipskamp.
- Caballo, V. E., Salazar, I. C., Iruiria, M. J., Arias, B., Hofmann, S. G., & CISO-A Research Team. (2008). Social anxiety in 18 nations: Sex and age differences. *Behavioral Psychology/Psicología Conductual, 16*, 163-187.
- Chorpita, B. F., Yim, L., Moffitt, C., Umemoto, L. A., & Francis, S. E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: A revised child anxiety and depression scale. *Behaviour Research and Therapy, 38*, 835-855.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Copeland, W. E., Angold, A., Shanaham, L., & Costello, E. J. (2014). Longitudinal patterns of anxiety from childhood to adulthood: The Great Smoky Mountains Study. *Journal of the American Academy of Child and Adolescent Psychiatry, 53*, 21-33.
- Diemer, J., Muhlberger, A., Pauli, P., & Zwanzger, P. (2014). Virtual reality exposure in anxiety disorders: Impacts on psychophysiological reactivity. *The World Journal of Biological Psychiatry, 15*, 427-442.
- Donnelly, A., Fitzgerald, A., Shevlin, M., & Dooley, B. (2018). Investigating the psychometric properties of the revised child anxiety and depression scale (RCADS) in a non-clinical sample of Irish adolescents. *Journal of Mental Health*. doi:10.1080/09638237.2018.1437604
- Ellis, L. K., & Rothbart, M. K. (2001, April). *Revision of the Early Adolescent Temperament Questionnaire*. Communication presented at the Biennial Meeting of the Society for Research in Child Development, Minneapolis. <http://www.bowdoin.edu/~sputnam/rothbart-temperament-questionnaires/pdf/lesa-ellis-srcd-poster-reprint.pdf>
- Fonseca-Pedrero, E., Fernández, L., Paino, M., Lemos-Giraldez, S., & Muñiz, J. (2013). Development of a computerized adaptive test for schizotypy assessment. *PLoS ONE, 8*, e73201.
- García-López, L. J., Piqueras, J. A., Díaz-Castela, M. M., & Inglés, C. J. (2008). Social anxiety disorder in childhood and adolescents: Current trends, advances and future directions. *Behavioral Psychology/Psicología Conductual, 16*, 501-533.
- García-López, L. J., Salvador, M., & de los Reyes, A. (2016). Assessment of social anxiety in adolescents. In K. Ranta, A. M. La Greca, L. J. Garcia-López, & M. Marttunen (Eds.), *Social anxiety and phobia in adolescents* (pp. 121-150). Amsterdam: Springer.
- Green, B. F. (1991). Guidelines for computer testing. In T. B. Gutkin & J. C. Conoley (Eds.), *The computer and the decision making process. Bums-Nebraska symposium on measurement and testing* (pp. 245-273). Hillsdale, MI: Erlbaum.
- Hamann, C., Schultze-Lutter, F., & Tarokh, L. (2016). Web-based assessment of mental well-being in early adolescence: A reliability study. *Journal of Medical Internet Research, 18*, e138.
- Hirai, M., Vernon, L. L., Clum, G. A., & Skidmore, S. T. (2011). Psychometric properties and administration measurement invariance of social phobia symptom measures: Paper-pencil vs. internet administrations. *Journal of Psychopathology and Behavioral Assessment, 33*, 470.
- IBM (2011). IBM SPSS Statistics for Windows, Version 20.0 [Computer software]. Armonk, NY: IBM Corp.

- Ilioudi, S., Lazakidou, A. A., Glezakos, N., & Tsironi, M. (2012). Health-related virtual communities and social networking services. In A. A. Lazakidou (Ed.), *Virtual communities, social networks and collaboration* (pp. 1-13). New York, NY: Springer.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36, 409-426.
- Kessler, R. C., Avenevoli, S., Costello, E. J., Georgiades, K., Green, J. G., Gruber, M. J., He, J.-P., Koretz, D., McLaughlin, K. A., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Merikangas, K. R. (2012). Prevalence, persistence, and sociodemographic correlates of DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. *Archives of General Psychiatry*, 69, 372-380.
- Kosters, M. P., Chinapaw, M. J. M., Zwaanswijk, M., Van der Wal, M. F., & Koot, H. M. (2015). Structure, reliability, and validity of the revised child anxiety and depression scale (RCADS) in a multi-ethnic urban sample of Dutch children. *BMC Psychiatry*, 15, 132.
- Mary Rothbart's Temperament Lab (2015). *The Early Adolescent Temperament Questionnaire (EATQ-R)*. Resource document. <http://www.bowdoin.edu/~sputnam/rothbart-temperament-questionnaires/instrument-descriptions/early-adolescent-temperament.html>
- McDonough-Caplan, H., Klein, D. N., & Beauchaine, T. P. (2018). Comorbidity and continuity of depression and conduct problems from elementary school to adolescence. *Journal of Abnormal Psychology*, 127, 326-337.
- Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, 93, 568-592.
- Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care*, 44, S69-S77.
- Muris, P., & Meesters, C. (2009). Reactive and regulative temperament in youths: Psychometric evaluation of the Early Adolescent Temperament Questionnaire-Revised. *Journal of Psychopathology and Behavioral Assessment*, 31, 7-19.
- Nikou, S. A., & Economides, A. A. (2017). Mobile-based assessment: Investigating the factors that influence behavioral intention to use. *Computers and Education*, 109, 56-73.
- Park, B. K., & Calamaro, C. (2013). A systematic review of social networking sites: Innovative platforms for health research targeting adolescents and young adults. *Journal of Nursing Scholarship*, 45, 256-264.
- Patalay, P., Hayes, D., Deighton, J., & Wolpert, M. (2016). A comparison of paper and computer administered Strengths and Difficulties Questionnaire. *Journal of Psychopathology and Behavioral Assessment*, 38, 242-250.
- Pineda, D., Martín-Vivar, M., Sandín, B., & Piqueras, J. A. (2018). Factorial invariance and norms of the 30-item shortened-version of the Revised Child Anxiety and Depression Scale (RCADS-30). *Psicothema*, 30, 232-237.
- Piqueras, J. A., Martín-Vivar, M., Sandín, B., San Luis, C., & Pineda, D. (2017). The Revised Child Anxiety and Depression Scale: A systematic review and reliability generalization meta-analysis. *Journal of Affective Disorders*, 218, 153-169.
- Piqueras, J. A., Pineda, D., Martín-Vivar, M., & Sandín, B. (2017). Confirmatory factor analysis and psychometric properties of the Revised Child Anxiety and Depression Scale-30 (RCADS-30) in clinical and non-clinical samples. *Revista de Psicopatología y Psicología Clínica*, 22, 183-196.
- Putnam, S. P., & Stifter, C. A. (2008). Reactivity and regulation: The impact of Mary Rothbart on the study of temperament. *Infant and Child Development*, 17, 311-320.
- Quero, S., Nebot, S., Rasal, P., Breton-Lopez, J., Baños, R. M., & Botella, C. (2014). Information and communication technologies for the treatment of small animal phobia in children. *Behavioral Psychology/Psicología Conductual*, 22, 255-274.

- Richardson, C. G., Johnson, J. L., Ratner, P. A., & Zumbo, B. D. (2009). The influence of web-versus paper-based formats on the assessment of tobacco dependence: Evaluating the measurement invariance of the dimensions of tobacco dependence scale. *Substance Abuse: Research and Treatment, 3*, 1-14.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*. doi: 10.18637/jss.v048.i02
- Rothbart, M. K. (2007). Temperament, development, and personality. *Current Directions in Psychological Science, 16*, 207-212.
- Rothbart, M. K., & Gartstein, M. A. (2008). Temperament. In M. M. Haith & J. B. Benson (Eds.), *Encyclopedia of Infant and Early Childhood Development* (pp. 318-332). Amsterdam: Elsevier.
- Sandín, B., Valiente, R. M., & Chorot, P. (2009). RCADS: Assessment of anxiety and depressive disorder symptoms in children and adolescents. *Revista de Psicopatología y Psicología Clínica, 14*, 193-206.
- Steele, R. M., Mummery, W. K., & Dwyer, T. (2009). A comparison of face-to-face or internet-delivered physical activity intervention on targeted determinants. *Health Education and Behavior, 36*, 1051-1064.
- Stevanovic, D., Bagheri, Z., Atilola, O., Vostanis, P., Stupar, D., Moreira, P., Franic, T., Davidovic, N., Knez, R., Nikšić, A., Dodig-Čurković, K., Avicenna, M., Noor, I. M., Nussbaum, L., Deljkovic, A., Thabet, A. A., Petrov, P., Ubalde, D., Monteiro, L. A., & Ribas, R. (2017). Cross-cultural measurement invariance of the Revised Child Anxiety and Depression Scale across 11 world-wide societies. *Epidemiology and Psychiatric Sciences, 26*, 430-440.
- Thompson, M. S., & Green, S. B. (2006). Evaluating between-group differences in latent means. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course*. London: Information Age.
- Van Ballegooijen, W., Riper, H., Cuijpers, P., Van Oppen, P., & Smit, J. H. (2016). Validation of online psychometric instruments for common mental health disorders: A systematic review. *BMC Psychiatry, 16*, 45.
- Weigold, A., Weigold, I., & Russell, E. J. (2013). Examination of the equivalence of self-report survey-based paper-and-pencil and internet data collection methods. *Psychological Methods, 18*, 53-70.
- Wyrick, D. L., & Bond, L. (2011). Reducing sensitive survey response bias in research on adolescents: A comparison of web-based and paper-and-pencil administration. *American Journal of Health Promotion, 25*, 349-352.

RECEIVED: January 9, 2018

ACCEPTED: June 20, 2018