A Brief Measure of Problematic Smartphone Use Among Youth:

Psychometric Assessment Using Item Response Theory

Scott I. Donaldson[a]

David Strong[ab]

Shu-Hong Zhu[ab]

[a]Moores Cancer Center, University of California San Diego, La Jolla, California, United States of America, [b]Department of Family Medicine and Public Health, University of California San Diego, La Jolla, California, United States of America.

sidonaldson@ucsd.edu

dstrong@ucsd.edu

szhu@ucsd.edu

Address correspondence to: Shu-Hong Zhu, Ph.D., University of California, San Diego, Cancer Center MC-0905, 9500 Gilman Drive, La Jolla, CA 92093-0905, United States of America. Telephone: 858-300-1056. Fax: 858-300-1099. E-mail: szhu@ucsd.edu.

Compliance with Ethical Standards

Conflict of Interest: All the authors declare no conflicts of interests.

Ethical Approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

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

Funding

The 2017-2018 California Student Tobacco Survey was supported by a contract from the California Department of Public Health (#16-10109). The interpretation of the results is solely the responsibility of the authors and does not necessarily represent the official views of the California Department of Public Health. The study funders did not play any role in conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

Abstract

The purpose of this study was to use item response theory to assess a brief measure of problematic smartphone use among youth, using the 2017-18 California Student Tobacco Survey (CSTS) collected from 119,981 students who own a smartphone across 256 high schools in California. An exploratory factor analysis supported two factors that represented problematic smartphone use and concurrent behavioral issues, explaining 47% of the variance. Item response modeling demonstrated good item discrimination for problematic smartphone use ($a > 1.15$) and valuable test information for respondents within two standard deviations of the sample mean. Students who reported a score of 3 (somewhat agree) or 4 (agree) on each problematic smartphone use item accounted for 22% ($n = 25{,}997$) of the student population who owned smartphones in our sample. Concurrent and criterion validity were found as problematic smartphone use significantly predicted smartphone use instead of sleep ($b = 0.35$, 95% *CI* [0.34, 0.36], $p < .05$), smartphone use instead of work ($b = 0.31$, 95% *CI* [0.30, 0.32], $p < .05$), depressive symptomatology ($OR = 1.34$, 95% CI [1.31, 1.37]), and loneliness ($b = 0.18$, 95% *CI* [0.16, 0.18], $p < .01$). Implications for screening and identifying appropriate cut-off criteria for problematic smartphone use among are discussed.

Keywords: Problematic smartphone use, Youth, Measurement, Psychometric, Item response theory

A Brief Measure of Problematic Smartphone Use Among Youth:
Psychometric Assessment Using Item Response Theory

1.0. Introduction

Multifunctional smartphones have transformed the youth psychosocial experience. From streamlining communication and information sharing to amalgamating social networks, or simply consuming spare time, the smartphone often plays an influential role in youth life. Although smartphones offer many benefits, youth can become overly reliant on their devices. Various studies have suggested that a significant proportion of adolescents may have developed problematic smartphone use (PSU) (Beasley et al., 2016; Tulane, Vaterlaus, & Bekert, 2018). For example, they spend time with their smartphones when they should be working or sleeping, which can have negative effects on their academic studies or their health (Demirci, Akgonul, & Akpinar, 2015; Thomee, Harenstam, & Hagberg, 2011). The over-utilization of smartphones has also been shown to predict negative emotions such as anxiety and depression (Bianchi & Phillips, 2005; Billieux, Maurage, Lopez-Fernandez, Kuss, & Griffiths, 2015; Choliz, 2016; Panova & Carbonell, 2018). A recent survey by Pew Research Center reports that the vast majority of American adolescents (95%) own a smartphone and nearly half worry that they spend too much time on it (Pew Research Center, 2019, August).

Research on PSU (or excessive smartphone use, smartphone addiction, compulsive smartphone use, and compensatory smartphone use; Kardefelt-Winther, 2014) has grown considerably in recent years. PSU has been compared to other behavioral disorders, such as gambling and internet addiction (Aljomaa, Mohammad, Albursan, Bakhiet, & Abduljabbar, 2016; Darcin et al., 2016; Demirci et al., 2015; Haug et al., 2015). While Griffiths (2013) warned against overpathologizing smartphone behaviors, PSU may share some behavioral symptom patterns with substance use disorders, such as conflict, motivational salience, and withdrawal. That is, intrapersonal, interpersonal, and social conflicts may arise from spending too much time on a smartphone (Griffiths, 2019; Kim, Lee, Lee, Nam, & Chung, 2014; Lin et al., 2014). Motivational salience can develop from overuse when people become preoccupied with using smartphones, such as constantly checking them for fear of missing conversations from friends on social media. Finally, physical discomfort (e.g., feeling irritable or restless) associated with being far from a smartphone could constitute withdrawal symptoms (Kwon, Kim, Cho, Yang, & Choi, 2013).

Several measures have been developed to assess PSU and related constructs. These include the Smartphone Addiction Scale (SAS-SV), Mobile Phone Involvement Questionnaire (MPIQ), Mobile Phone Problematic Use Scale (MPPUS), Problematic Mobile Phone Use Questionnaire (PMPUQ), and the Test of Mobile-Phone Dependence among others (Bianci & Phillips, 2005; Billieux, Van der Linden, & Rochat, 2008; Carbonell, Chamarro, Oberst, Rodrigo, & Prades, 2018; Chen et al., 2017; Choliz, 2012; Kwon et al., 2013; Matar & Jaalouk, 2017; Walsh, White, & Young, 2010). Studies have demonstrated that these measures of PSU predict sleep and work disturbances, and depressive symptomatology among other behavioral issues (Alhassan et al., 2018; Demirci et al., 2015; Elhai, Dvorak, Levine, & Hall, 2017).

1.1. The current study

The present study aims to develop a brief measure of PSU that can be used in population surveys to monitor adolescent behaviors. A brief measure is more likely to be adopted in population surveys than a full scale since population surveys often assess multiple behaviors, making it difficult to include a full scale for one particular behavior. Additionally, population surveys provide an excellent opportunity to test the general usefulness of a measure because they often recruit representative samples from the population. In the present case, we tested a brief scale of PSU with a large (>100,000) random sample of high school students in California who participated in a survey on tobacco use behaviors.

Item response theory (IRT) tested the psychometric properties of a brief measure of PSU. Item response theory offers several well-established methods to establish construct validity, estimate item characteristics, and quantify precision of PSU measurement in a diverse and large population of youth. To date, the majority of PSU measures were validated on university students and adults, often with international samples. This study focuses on youth. The brief scale chosen for this study consists of three items adapted from the 10-item SAS-SV by Kwon and colleagues (Kwon et al., 2013).

2. Method

2.1. A population survey of high school students in California

We used data from the 2017-2018 California Student Tobacco Survey (CSTS), amassing 256 high schools in California. The purpose of the CSTS is to assess the prevalence, the knowledge of, and attitudes toward cigarettes and other tobacco products. An online survey was administered during school hours that took 15–25 minutes to complete. For those schools with insufficient computer access, tablets and proctors from the University of California, San Diego (UCSD) were offered. Each school received a $500 gift card as an incentive to participate in the survey.

A two-stage cluster sampling design was used with school as the primary unit and classroom as the secondary unit. Public schools and non-sectarian schools were included in the survey. Schools that were special education, juvenile court, district/county community, continuation, online, and other forms of alternative education were excluded from the survey. A probability proportion to size (PPS) sampling approach with size determined by enrollment was used with 22 regions based on socioeconomic characteristics. Schools were then stratified within region with rural, African-American, and Tier 2 Tobacco-Use Prevention Education (TUPE) funded schools being oversampled to account for small sample size.

2.2. Measures

Participants responded *yes/no* or "*I prefer not to answer*" to the item, "Do you have a smartphone?" Sociodemographic variables included gender (male, female, other – I identify my gender in another way or I prefer not to answer), grade level (10th, 12th), and race/ethnicity (Non-Hispanic (NH) White, NH-Black, Hispanic, NH-Asian, NH-AI/AN (American Indian/Alaska Native), NH-NHOPI (Native Hawaiian and Other Pacific Islanders), NH-Other, NH-Multiple races).

We included three PSU items (items 5, 8, & 10 from the SAS-SV) using declarative statements, which were adapted from Kwon et al. (2013) and measured conflict, motivational salience, and withdrawal core components of addiction (Bianchi & Phillips, 2005; Csibi, Griffiths, Cook, Demetrovics, & Szabo, 2018; Griffiths, 2008; Kwon et al., 2013). Conflict was assessed with the item "I feel like my parent/legal guardian is always asking me to stop using my phone" on a 4-point Likert scale (1-Strongly Disagree, 4-Strongly Agree). The salience item was measured on a 4-point Likert scale and stated, "I feel like I have to constantly check my phone so I do not miss conversations from my friends on social media (e.g., Facebook, Instagram, or Twitter)." Lastly, we included a withdrawal item on a 4-point Likert scale (1-Never, 4-Always). The item stated, "I feel impatient or uncomfortable when I do not have my phone with me." Participants could elect not to answer any question. *Conflict*, *motivational salience*, and *withdrawal* were denoted SP (smartphone)_Parent, SP_Constant, and SP_Uncomfort, respectively.

Next, three items were selected from the CSTS to assess behavioral issues of PSU. These items included sleep and work disturbances, and using smartphones in a socially awkward situation. The CSTS asked about participants' sleep and work habits related to smartphone use on a 4-point Likert scale (1-Never, 4-Alway*s*). Those questions were "How often do you stay up at night using your smartphone, when you should be sleeping?" and "How often do you use your smartphone when you are supposed to be working?" We assessed smartphone use in socially awkward situations with one item on a 4-point Likert scale, ranging from 1 (Never) to 4 (*Always*). The item asked, "How often do you use (or pretend to use) your smartphone when you are in a socially awkward situation?" *Sleep*, *work*, and *smartphone use in a socially awkward situation* were denoted SP_Sleep, SP_Work, and SP_Soc_Awk, respectively.

Our final analyses measured loneliness with the item "A lot of times I feel lonely" on a 4-point Likert scale (1-Strongly Disagree, 4-Strongly Agree). Depressive symptomatology included a yes or no response option on this survey question, "In the last 12 months, did you ever feel sad and hopeless EVERYDAY for 2 weeks or more?"

2.3. Analytic Strategy

All analyzes were conducted in R (version 3.6.1, R Foundation for Statistical Computing, 2019) using the Mokken (Andries van der Ark, 2007), KernSmooth IRT (Mazza, Punzo, & McQuire, 2014), mirt (Chalmers, 2012), ltm (Rizopoulos, 2006), survey (Lumley, 2020), and effects (Chambers & Hastie, 1992) packages.

2.3.2 Dimensionality

We applied multidimensional IRT to conduct an expectation-maximization exploratory factor analysis (EFA) with oblique rotation using polychoric correlations. Factor extraction was guided by parallel analysis of randomly generated values, a scree plot, the requirement that modeled factors were supported by eigenvalues greater than one, patterns of strong factor loadings, and a match to theoretical expectations for the presence of more than one construct reflected within the selected item pool. Goodness of fit statistics included the root mean square error of approximation (RMSEA), Akaike information criterion (Akaike, 1974), Bayesian information criterion (BIC; Bentler & Bonett, 1980), and a chi-square statistic calculated as the difference in deviance (G^2). The purpose of the EFA procedure was to examine the dimensionality of our PSU scale and degree of separation from concurrent items reflecting behavioral issues.

2.3.3. Nonparametric item response theory

We used a Nonparametric IRT (NIRT) and Mokken Scale Analysis to assess six polytomous survey items (Mokken, 1971). Mokken Scale Analysis is useful for exploring the psychometric properties of data before implementing a parametric IRT model (Meijer & Banke, 2004). Four assumptions were tested in our NIRT model: unidimensionality, local independence, latent monotonicity, and nonintersection. First, the scalability coefficient H was used for pairs of items and the total set of items in the test. Mokken (1971, p. 185) suggested that H =.4 denotes a weak scale, .4 < H <.5 denotes a medium scale, and H > .5 denotes a strong unidimensional scale. Coefficient alpha and lamba also assessed internal consistency. We then interrogated unidimensionality with the automated item selection procedure, which calculates inter-item covariances and the relationship between items and the latent trait (Meijer & Banke, 2004). Latent monotonicity for each item was examined using a visual plot of item step response function by rest score group. The rest score group is achieved by summing the overall score minus the score on each item (Junker, 1993).

Finally, we used Kernel Smoothing techniques to fit nonparametric option characteristic curves to visually examine monotonicity given variability along a single latent construct (Mazza, Punzo, & McGuire, 2014). We did not assess clinical thresholds for diagnosing PSU in our sample. Future research will need to determine the utility and appropriateness of cut-off scores (i.e., youth who are addicted versus not addicted) using our measure in an independent, representative sample of youth.

2.3.4. Parametric item response theory

After investigating our NIRT models, a full information graded response model with sample weights was used to estimate item discrimination and ability on the latent trait (Santor & Ramsay, 1998). Sample weights and school-level stratification data accounted for the effect of school-level variation when estimating graded response models. We performed a residual analysis at the item level to evaluate local independence, inspecting for plots between -2 and 2 (Zanon, Hutz, Yoo, & Hambleton, 2016). We also estimated the test information function, an estimate of item measurement precision across a range of observed scores on the latent trait. Finally, for each item we estimated item and option characteristics curves and person-level estimates using expected a posterior scores from the model.

2.3.5. Regression

Our final analyses fit survey-weighted generalized linear and logistic regression models to predict smartphone use instead of sleep and work, depressive symptomatology, and loneliness. We included sample weights and school-level stratification data to account for the effect of school-level variation on our regression findings. Working likelihood ratio (Rao-Scott) tests were used to reveal significant independent variables (Rao & Scott, 1984).

3. Results

3.1. Survey demographics

A total of 95.5% (119,981/125,546) of 10th and 12th grade students from the CSTS reported smartphone ownership. Table 1 presents the demographic characteristics of those students who own and those students who do not own smartphones. The demographics for the two groups were very similar for the most part. One difference was observed between students whose gender was considered "other" for purposes of this study, meaning they selected "I identify my gender in another way" or "I prefer not to answer." Students of other gender comprised a higher percentage of the total sample of students (7.5%) who did not own smartphones compared to other gender students who did (2.4%). In addition, a greater proportion of 10th grade students compared to 12th grade students did not own a smartphone.

The sample of smartphone users that we included was represented by female (48.4%), male (43.9%), and other (2.4%). Approximately 5.3% of students declined to state their gender. The primary composition of race/ethnicity was 19.9% Non-Hispanic White, 49.1% Hispanic, and 11.6% Asian. We only included high school students who completed the CSTS, which included 54.0% 10th graders and 46.0% 12th graders. Full information maximum likelihood was used for the IRT analyses, resulting in a final sample of 108,224 students. Listwise deletion was used to handle missing data on behavioral issues for the regression analyses, resulting in a suitable sample of 96,701 high school students.

Table 1
Sociodemographic characteristics of study participants.

	Students who own smartphones	Students who do not own smartphones
Demographic variable	Weighted percent and 95% confidence interval	Weighted percent and 95% confidence interval
Gender		
Female	48.4 (48.1, 48.7)	36.8 (35.5, 38.1)
Male	43.9 (42.8, 44.1)	43.8 (42.5, 45.1)
Other	2.4 (2.3, 2.5)	7.5 (6.8, 8.2)
Decline to state	5.3 (5.1, 5.4)	11.9 (11.0, 12.8)
Race/Ethnicity		
NH-White	19.9 (19.7, 20.2)	11.8 (10.9, 12.6)
NH-Black	2.5 (2.4, 2.6)	3.9 (3.4, 4.4)
Hispanic	49.1 (48.8, 49.4)	51.9 (50.6, 53.3)
NH-Asian	11.6 (11.4, 11.8)	7.6 (6.9, 8.3)
NH-AI/AN	0.27 (0.25, 0.31)	0.9 (.06, 1.1)
NH-NHOPI	0.60 (0.56, 0.65)	1.4 (1.0, 1.7)
NH-Other	1.6 (1.5, 1.7)	2.4 (2.0, 2.8)
NH-Multiple race	8.8 (8.7, 9.0)	7.6 (6.9, 8.3)
Grade		
10	54.0 (53.7, 54.3)	62.9 (61.6, 64.3)
12	46.0 (45.7, 46.3)	37.1 (35.8, 38.4)

Note. Other = I identify my gender in another way or I prefer not to answer; NH = Non-Hispanic, AI/AN = American Indian or Alaska Native, NHOPI = Native Hawaiian and Other Pacific Islander; N = 125,546.

The average PSU score was 2.21 (SD = 0.79). Skewness and kurtosis scores were within the normal range. Tertile scores reflecting low, medium, and high PSU scores were 1.6 (33rd percentile), 2.7 (66th percentile), and 4.0 (99th percentile), respectively. Students who reported a score of 3 (somewhat agree) or 4 (agree) on each PSU item accounted for 24% (n = 25,997) of the student population who owned smartphones. Finally, students who reported two or more 4s (agree) on the PSU measure represented 6.2% (n = 6,695) of the student population who owned smartphones. Figure 1 shows the distribution of problematic smartphone use scores.

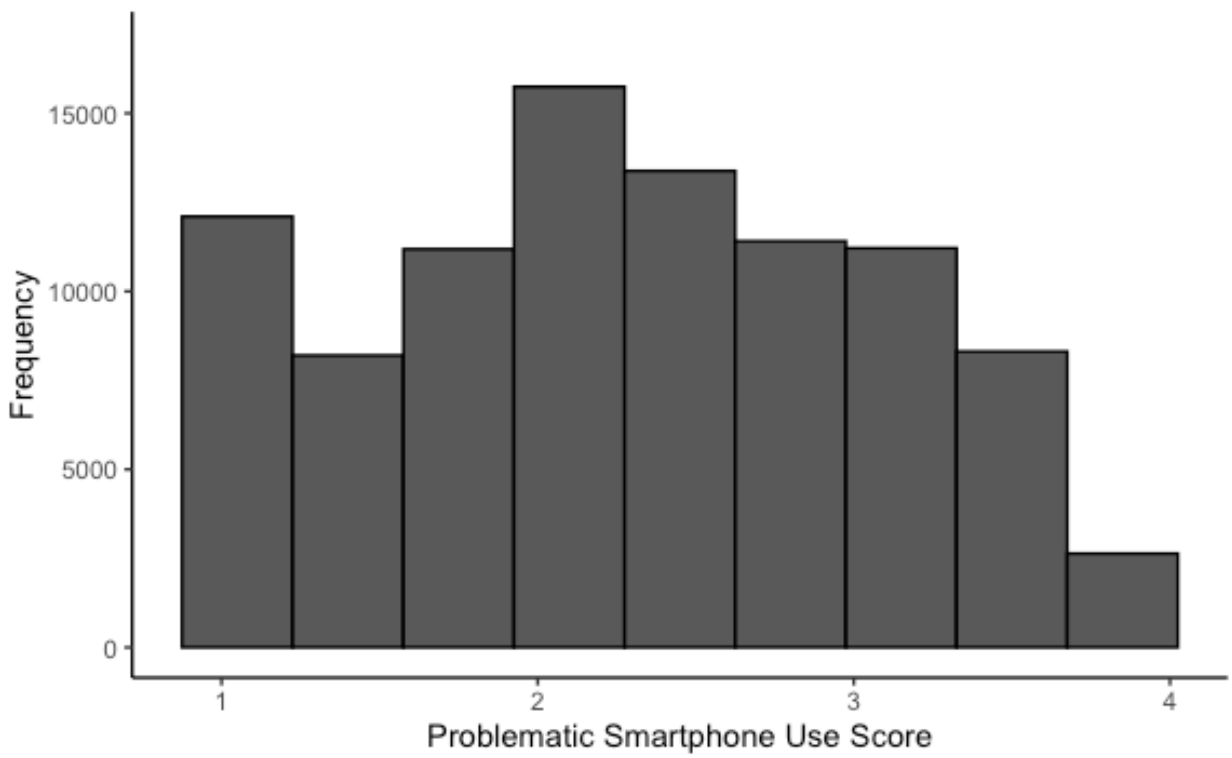

Figure 1 Distribution of problematic smartphone use.

3.2. Dimensionality

Factor analysis was used to explore the underlying dimensions of PSU and behavioral issues. Inter-item correlations, item variances, item means, and coefficients of internal subscales were examined. Parallel analysis was conducted to help determine the best number of factors to retain (Horn, 1965), and a scree plot with simulated and real data suggested approximately two factors. Using survey weights, we compared a unidimensional and two-dimensional EFA factor structure using a graded response model. Model fit comparisons with AIC and log likelihood estimates revealed the multidimensional model fit the data best ($\Delta\chi^2$ =10292.18, df = 5, ΔAIC = 10282, ΔBIC = 10423). The EFA revealed a two-factor solution with acceptable factor loadings (> .40) and small cross-loadings (<.15; Devellis, 2012; Young & Pearce, 2013). The two-factor solution demonstrated adequate fit ($G^2(4066) = 23520.39$, $p < .01$; *RMSEA* = .008; *BIC* = 1227471; *AIC* = 1227200) and accounted for 47% of the variance in item responses (28% for PSU and 19% for behavioral issues). The two factors reflecting PSU and behavioral issues showed a strong positive correlation ($r = .59$). Items on the behavioral issues factor showed communality estimates below .50, which Costello and Osborne (2005) suggest is a minimum cutoff. As such, scalability of response patterns for these items were further explored alongside items from each separated factor in the Mokken Scale Analysis. Please see Table 2 for full information exploratory factor loadings, communalities, and item parameters.

Table 2
Full information for exploratory factor loadings, communalities, and item parameters for PSU and behavioral issues.

		Rotated Factor Loadings and Item Parameters						
Item	M (SD)	λ^a	λ^b	h^2	a	b_1	b_2	b_3
SP_Uncomfort[a]	2.4 (1.0)	0.76		0.61	2.05	-0.92	0.07	1.33
SP_Constant[a]	2.0 (.98)	0.85		0.70	3.21	-0.39	0.51	1.50
SP_Parent[a]	2.1 (1.0)	0.50		0.28	1.09	-0.71	0.60	2.07
SP_Sleep[b]	2.7 (.92)		0.61	0.40				
SP_Work[b]	2.2 (.84)		0.72	0.48				
SP_Soc_Awk[c]	2.7 (1.0)		0.41	0.26				

Note. ^{a} = Problematic Smartphone Use; ^{b} = Behavioral Issues; *c* = SP_Soc_Awk was excluded from behavioral issues after EFA; λ = factor loading, h^2 = Communality; *a* = Discrimination; *b* = Threshold.

3.3. Item response modeling

Mokken Scale Analysis of the PSU scale items revealed a medium item-pair scalability coefficient ($H = 0.46$) and a weak item-pair scalability coefficient ($H = 0.36$) for items reflecting behavioral issues (Mokken, 1971). All PSU items were retained at this stage. Due to the low communality and item-pair scalability coefficient, we decided to drop SP_Soc_Awk from behavioral issues, which increased H into the medium range ($H = 0.46$) and reduced the set to two items. For all further analyses, we used two indices of behavioral issues: a) smartphone use instead of sleep and b) smartphone use instead of work. An automated item selection algorithm and plots further supported a PSU scale and revealed monotonic relationships with the rest score group. Examination of OCC plots for PSU suggested that all three items provided distinct separation across the four response options and were suitable for parametric analyses (see Figure 2).

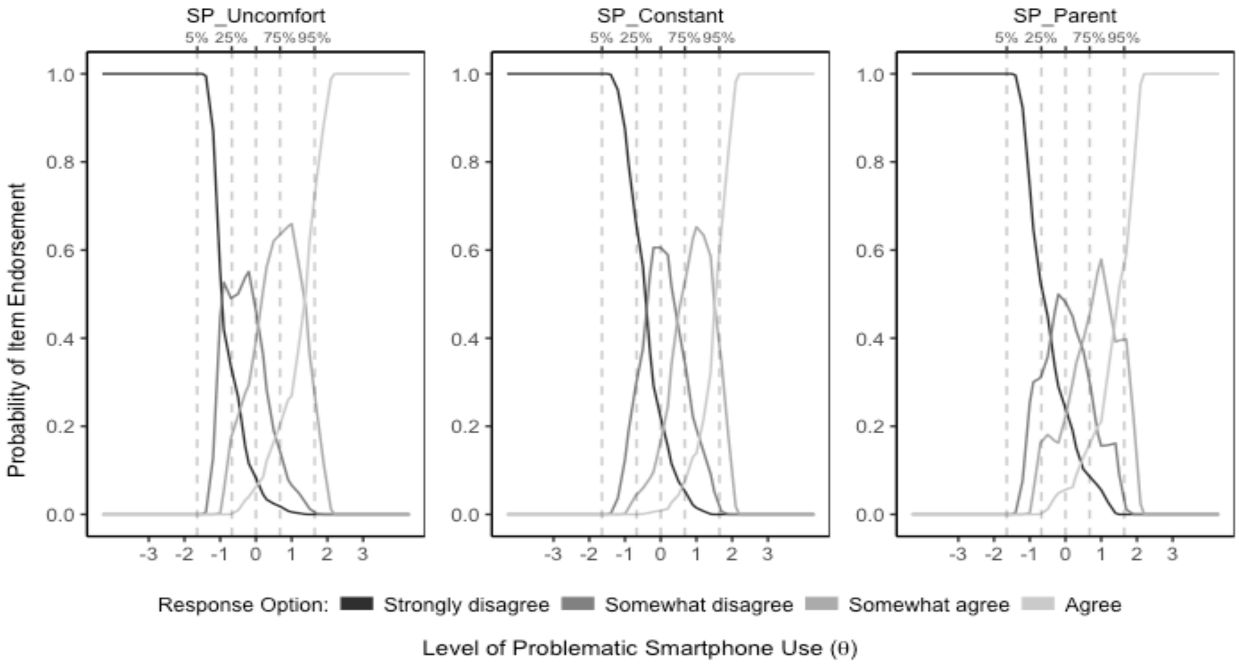

Figure 2 Kernel smoothed item response models for problematic smartphone use.

Parametric IRT analyses found that SP_Uncomfort and SP_Constant had higher item discrimination across PSU than the SP_Parent item discrimination (Baker, 1985). An empirical plot of residuals versus predicted scores revealed response categories 2 (*Sometimes*) and 3 (*Often*) did not fit well in the given model. Given the weaker fit of SP_Parent, we collapsed response categories 2 and 3, which improved coefficient H to 0.52 (i.e., medium to strong) and improved item residuals in concordance with the item characteristic curve. A nested chi-square comparison revealed a significantly better fit ($\Delta\chi^2 = 241660.3$, $df = 42$, $\Delta AIC = 241668.3$ $\Delta BIC = 241705.7$) with the collapsed SP_Parent item. As demonstrated in Figure 3, most of the test information was between one standard deviation above and below the mean, suggesting that the scale was designed for the majority of respondents with PSU observed in this population.

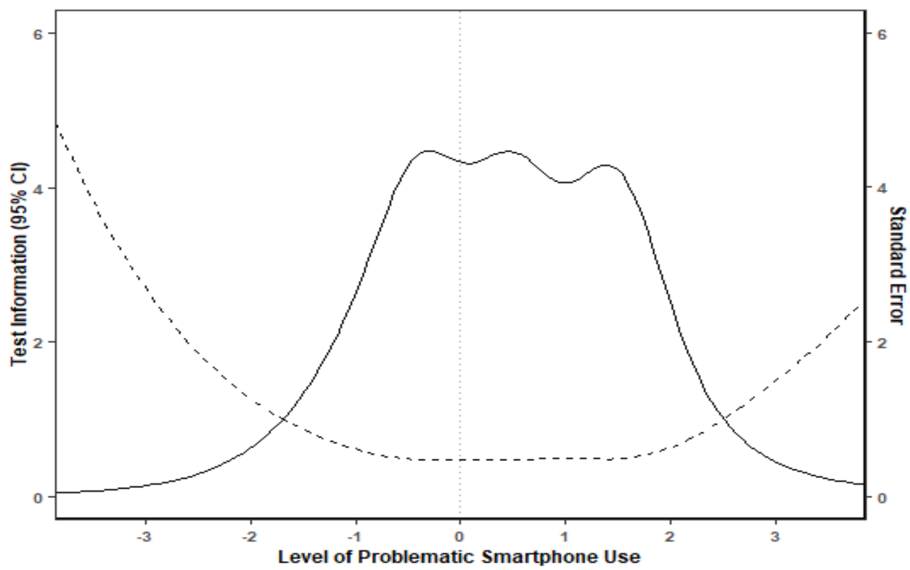

Figure 3 Test information function (solid line) and standard errors (broken line) for problematic smartphone use.

3.4. Reliability, concurrent, and criterion validity

Cronbach's α (α = .69; Cronbach, 1951) was in the acceptable range for the PSU scale. Multiple regression analyses with survey weights were then run to associate sleep disturbance and work disturbance with PSU, controlling for gender, race, and grade.

The linear regression model of sleep disturbance was statistically significant and explained 10% of this item's variance. Controlling for demographic variables, PSU was significantly associated with sleep disturbance, $b = 0.35$, 95% *CI* [0.34, 0.36], $p < .05$ (see Table 3). Females were significantly more likely than males to experience sleep disturbances ($b = 0.03$, 95% *CI* [0.02, 0.05], $p < .05$), as well as other gender identification ($b = 0.07$, 95% *CI* [0.01, 0.13], $p = 0.28$) compared to males. Tenth grade students ($b = 0.15$, 95% *CI* [0.14, 0.17]), $p < .05$) were more likely to experience sleep disturbances related to smartphone use than 12th grade students. Compared to Whites, NH-Black ($b = 0.18$, 95% *CI* [0.13, 0.22], $p < .05$), Hispanic ($b = 0.04$, 95% *CI* [0.02, 0.06], $p < .05$), NH-Asian ($b = 0.07$, 95% *CI* [0.04, 0.09], $p < .05$), and NH-NHOPI ($b = 0.19$, 95% *CI* [0.11, 0.27], $p < .05$), and NH-Multiple Race ($b = 0.07$, 95% *CI* [0.04, 0.10], $p < .05$) were significantly more likely to experience sleep disturbances. No significant effects of smartphone use on sleep were found for the comparison between Whites, NH-AI/AN, and NH-Other (see Table 3).

Table 3 also shows that the relationships between PSU ($b = 0.31$, 95% *CI* [0.30, 0.32), $p < .05$) and work disturbances were statistically significant, explaining 11% of this item's variance. In terms of gender, females ($b = 0.12$, 95% *CI* [0.11, 0.14], $p < .05$) and other ($b = 0.15$, 95% *CI* [0.09, 0.20], $p < .05$) were significantly more likely than males to experience work disturbances. No work disturbances were observed between 10th and 12th grade students. Compared to Whites, NH-Black ($b = 0.14$, 95% *CI* [0.10, 0.19], $p < .05$), NH-Asian ($b = 0.17$, 95% *CI* [0.15, 0.19], $p < .05$), NH-NHOPI ($b = 0.06$, 95% *CI* [-0.02, 0.14], $p < .05$), and NH-Multiple Race ($b = 0.06$, 95% *CI* [0.04, 0.09], $p < .05$) were significantly more likely to experience work disturbances. On the other hand, Hispanic (*b* = -0.08, 95% *CI* [-0.10, -0.06], $p < .05$), NH-AI/AN (*b* = -0.11, 95% *CI* [-0.23, 0.02], $p < .05$), and NH-Other (*b* = -0.06, 95% *CI* [-0.12, -0.01], $p < .05$) were significantly less likely to experience work disturbances compared to Whites.

Our final analyses modelled the relationship between PSU, loneliness, and depressive symptomatology. Controlling for demographic characteristics, a survey-weighted generalized linear regression demonstrated that PSU was significantly associated with loneliness (*b* = 0.18, 95% *CI* [0.16, 0.18), $p < .01$, $R^2 = 0.05$), and significantly associated with depressive symptomatology (*OR* = 1.34, 95% CI [1.31, 1.37], R^2 (McFadden) = 0.04).

Table 3
Survey-weighted generalized linear regression predicting sleep and work disturbances among smartphone owners (*n* = 96,072)

	Sleep		Work	
Variable	*b*	95% CI	*b*	95% CI
Intercept	1.78^{*}	(1.75, 1.80)	1.48^{**}	(1.46, 1.51)
Problematic Smartphone Use	0.35^{*}	(0.34, 0.36)	0.31^{**}	(0.30, 0.32)
Gender				
Female vs. Male	0.03^{*}	(0.02, 0.05)	0.12^{**}	(0.11, 0.14)
Other vs. Male	0.07^{*}	(0.01, 0.13)	0.15^{**}	(0.09, 0.20)
Grade Level				
10th vs. 12th	0.15^{*}	(0.14, 0.17)	0.00	(-0.01, 0.01)
Race				
NH-Black vs. White	0.18^{*}	(0.13, 0.22)	0.14^{**}	(0.10, 0.19)
Hispanic vs. White	0.04^{*}	(0.02, 0.06)	-0.08^{**}	(-0.10, -0.06)
NH-Asian vs. White	0.07^{*}	(0.04, 0.09)	0.17^{**}	(0.15, 0.19)
NH-AI/AN vs. White	-0.01	(-0.13, 0.11)	-0.11	(-0.23, 0.02)
NH-NHOPI vs. White	0.19^{*}	(0.11, 0.27)	0.06*	(-0.02, 0.14)
NH-Other vs. White	0.04	(-0.03, 0.11)	-0.06^{**}	(-0.12, -0.01)
NH-Multiple Race vs. White	0.07^{*}	(0.04, 0.10)	0.06^{**}	(0.04, 0.09)

Note. $^{*} = p < .05$; Sleep ($R^2 = 0.10$); work ($R^2 = 0.11$).

4. Discussion

This study provided an IRT analysis of a brief measure of problematic smartphone use using a population-level sample of California youth. Taken together, item response modeling described a set of coherent items reflecting problematic smartphone behaviors that could be scaled separately from sleep and work-related issues of problematic smartphone use. The problematic smartphone use scale items demonstrated acceptable item discrimination, thresholds, and, when summed, provided test information within two standard deviations of the normally distributed problematic smartphone use scores. Descriptive analyses showed that nearly 22% of the student population who owned smartphones in our sample at least somewhat agreed that they use their smartphones in problematic ways. We also found support for relationships between problematic smartphone use scores and concurrent behavioral issues, including sleep and work disturbances related to smartphone use, as well as loneliness and depressive symptomatology. With initial validation of an instrument in a representative sample of youth in California, future research can establish meaningful threshold criteria for screening this technology-related behavioral problem and explore potential associations with health consequences among youth.

Several efforts to validate problematic smartphone use measures were initiated with university students and adults internationally, including samples from South Korea, Australia, Spain, and Lebanon to name a few. Our findings advance measurement precision in this literature using a large, diverse sample of high school students in California. Our psychometric analyses provide item level information of behaviors relevant to youth in the United States and produced a normally distributed measure of problematic smartphone use scores. This empirical contribution using methods based in IRT should be viewed as a complement to existing scale validation studies, and may encourage further research to link efforts to develop instruments reflecting behavioral problems and extend measurement across age ranges using innovative measurement techniques.

It is well documented that over-reliance on smartphones may in tandem interfere with sleep or work, and lead to relationship difficulties (Darcin et al., 2016; Demirci et al., 2015; Gao et al., 2016; Kim, Kim, & Jee, 2015; Kuss & Griffiths, 2011; Tan, Pamuk, & Donder, 2013). For example, Irwin et al. (1996) found that most teens use their cell phones one hour before sleep, which is linked to a number of negative outcomes, including sleep dysregulation and daytime dysfunction (Demirci et al., 2015; Van den Bulck, 2007). Thomee et al. (2011) found that high frequency of smartphone use was a risk factor for developing sleep disturbances in men and depressive symptomatology in both men and women. Our study found that females/other and students from certain disparity populations (e.g., African Americans) were predisposed to higher levels of sleep and work disturbances. Thus, these students may be more vulnerable to the consequences of problematic smartphone use.

We also found significant relationships between problematic smartphone use and depressive symptomatology and loneliness. Demirci et al. (2015) demonstrated a link between problematic smartphone use and behavioral issues, such as depression and anxiety. One explanation is that increases in screen time and simultaneous decreases in non-screen activities (e.g., social interaction) explain proliferating depression rates among U.S. youth (Twenge, Martin, & Campbell, 2018). Other evidence suggests a bidirectional relationship between depressive symptoms and problematic smartphone use. Depressive individuals may escape negative emotions by using their smartphone, which may elicit more depression, irritability, and stress (Elhai et al., 2018). Likewise, excessive smartphone use may be an effort to overcome feelings of loneliness by seeking out opportunities for socialization (Darcin et al., 2015); however, difficulties engaging social contacts online could potentially reinforce feelings of loneliness.

This study has limitations that deserve some attention. First, the purpose of the CSTS is to assess tobacco use in youth and includes only a few items about smartphone behaviors. As such, there were restrictions on the number of items available for psychometric validation and criterion testing (i.e., single item measures). Future studies would benefit from comparing these results to other validated scales in the literature to better understand the potential negative affect and interpersonal disconnectedness that may be associated with problematic smartphone use. Second, the CSTS is a cross-sectional survey using self-reported data from high school students. Thus, there was no opportunity to develop validation criteria for cut-off scores. Future studies could assess the role of differential item functioning between youth and adults, as well as investigate the age at which behavioral problems emerge and their influence on the developmental trajectories of problematic smartphone use relative to behavioral issues.

5.0 Conclusion

This study using IRT found support for psychometric properties of a brief measure of problematic smartphone use. The item level information of this measure advances measurement precision on problematic smartphone use and predicts co-occurring behavioral issues, such as depressive symptomatology and sleep and work disturbances. Future research should further examine this smartphone behavior among youth in order to develop screening tools and interventions that can effectively ameliorate this emergent behavior.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Caski (Eds.), *Proceedings of the second international symposium on information theory* (pp. 267–281). Budapest: Akademiai Kiado.

Alhassan, A. A., Alqadhib, E. M., Taha, N. W., Alahmari, R. A., Salam, M., & Almutairi, A. F. (2018). The relationship between addiction to smartphone usage and depression among adults: a cross sectional study. *Bmc Psychiatry*, *18*(1).

Aljomaa, S. S., Al.Qudah, M. F., Albursan, I. S., Bakhiet, S. F., & Abduljabbar, A. S. (2016). Smartphone addiction among university students in the light of some variables. *Computers in Human Behavior*, *61*, 155–164.

Baker, F. B. (1985). *The basics of item response theory* (1st ed.). Heinemann.

Beasley, R. E., McMain, J. T., Millard, M. D., Pasley, D. A., Western, M. J. (2016, October). The Effects of College Student Smartphone Use on Academic Distraction and Dishonesty. Paper presented at the Consortium for Computing Sciences in Colleges, Evansville, IN.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588–606.

Bernaards, C. A., & Jennrich, R. I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement*, *65*(5), 676–696.

Bianchi, A., & Phillips, J. G. (2005). Psychological predictors of problem mobile phone use. *Cyberpsychol Behav*, *8*(1), 39-51.

Billieux, J. (2012). Problematic use of the mobile phone: A literature review and a pathways model. *Current Psychiatry Reviews, 8*(4), 299–307.

Billieux, J., Maurage, P., Lopez-Fernandez, O., Kuss, D. J., & Griffiths, M. D. (2015). Can disordered mobile phone use be considered a behavioral addiction? An update on current evidence and a comprehensive model for future research. *Current Addiction Reports*, *2*(2), 156–162.

Billieux, J., Van Der Linden, M., & Rochat, L. (2008). The role of impulsivity in actual and problematic use of the mobile phone. *Applied Cognitive Psychology, 22*(9), 1195–1210.

Carbonell, X., Chamarro, A., Oberst, U., Rodrigo, B., & Prades, M. (2018). Problematic use of the internet and smartphones in university students: 2006-2017. *International Journal of Environmental Research and Public Health*, *15*(3).

Chalmers, R., P. (2012). Mirt: A multidimensional item response theory package for the r environment. *Journal of Statistical Software, 48*(6), 1-29.

Chambers, J. M., & Hastie, T. J. (1992). *Statistical models.* Wadsworth & Brooks/Cole Advanced Books & Software.

Chen, B., Liu, F., Ding S., Ying, X., Wang, L., & Wen, Y. (2017) Gender differences in factors associated with smartphone addiction: A cross-sectional study among medical college students. *BMC Psychiatry*, *17*(1).

Choi, S.-W., Kim, D.-J., Choi, J.-S., Ahn, H., Choi, E.-J., Song, W.-Y., … Youn, H. (2015). Comparison of risk and protective factors associated with smartphone addiction and internet addiction. *Journal of Behavioral Addictions*, *4*(4), 308–314.

Chóliz, Mariano. (2016). The challenge of online gambling: The effect of legalization on the increase in online gambling addiction. *Journal of Gambling Studies*, *32*(2).

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*(7), 1–9.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.

Csibi, S., Griffiths, M. D., Cook, B., Demetrovics, Z., & Szabo, A. (2018). The psychometric properties of the smartphone application-based addiction scale (SABAS). *International Journal of Mental Health and Addiction*, *16*(2), 393–403.

Darcin, A., Kose, S., Noyan, C. O., Nurmedov, S., Yılmaz, O., & Dilbaz, N. (2016). Smartphone addiction and its relationship with social anxiety and loneliness. *Behaviour & Information Technology*, *35*(7).

Demirci, K., Akgönül Mehmet, & Akpinar, A. (2015). Relationship of smartphone use severity with sleep quality, depression, and anxiety in university students. *Journal of Behavioral Addictions*, *4*(2), 85–92.

Elhai, J. D., Dvorak, R. D., Levine, J. C., & Hall, B. J. (2017). Problematic smartphone use: A conceptual overview and systematic review of relations with anxiety and depression psychopathology. *Journal of Affective Disorders*, *207*, 251–259.

Gao, Y., Ang, L., Tingshao, Z., Xiaoqian, L., & Xingyun, L. (2016). How smartphone usage correlates with social anxiety and loneliness. *Peerj*, *4*.

Griffiths, M. D. (2008). Videogame addiction: Fact or fiction? In T. Willoughby & E. Wood (Eds.), *Children's learning in a digital world* (pp. 85–103). Blackwell Publishing.

Griffiths, M. D. (2013). Is ‘loss of control’ always a consequence of addiction? *Frontiers in Psychiatry*.

Griffiths, M. D. (2019, March). *Smartphone addiction, revisited.* Retrieved from https://www.psychologytoday.com/us/blog/in-excess/201903/smartphone-addiction-revisited

Gutiérrez J. D. S., Rodríguez, F. F., & Rubio, G. (2016). Cell-phone addiction: a review. *Frontiers in Psychiatry*, *7*, 175–175.

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, *10*(4), 255–282.

Haug, S., Castro, R. P., Kwon, M., Filler, A., Kowatsch, T., & Schaub, M. P. (2015). Smartphone use and smartphone addiction among young people in Switzerland. *Journal of Behavioral Addictions*, *4*(4), 299–307.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185.

Karadağ E, Tosuntaş ŞB, Erzen, E., Duru, P., Bostan, N., Şahin BM, … Babadağ B. (2015). Determinants of phubbing, which is the sum of many virtual addictions: A structural equation model. *Journal of Behavioral Addictions*, *4*(2), 60–74.

Kardefelt-Winther, D. (2014). A conceptual and methodological critique of internet addiction research: towards a model of compensatory internet use. *Computers in Human Behavior*, *31*(1), 351–354.

Kim, S. E., Kim, J. W., & Jee, Y. S. (2015). Relationship between smartphone addiction and physical activity in Chinese international students in Korea. *Journal of Behavioral Addictions*, *4*(3), 200–5.

Kim, D., Lee, Y., Lee, J., Nam, J. E. K., Chung, Y., & Bruce, A. (2014). Development of Korean smartphone addiction proneness scale for youth. *Plos One*, *9*(5), 97920.

Kwon, M., Kim, D. J., Cho, H., Yang, S., & Choi, D. S. (2013). The smartphone addiction scale: development and validation of a short version for adolescents. *Plos One*, *8*(12).

Lumley, T (2020). *Survey: Analysis of complex survey samples*. Retrieved from https://cran.r-project.org/web/packages/survey/citation.html

Matar, B. J., & Jaalouk, D. (2017). Depression, anxiety, and smartphone addiction in university students: A cross sectional study. *Plos One*, *12*(8), 0182239.

Mazza, A., Punzo, A., & Mc-Guire, B. (2014). Kernsmoothirt: An r package for kernel smoothing in item response theory. *Journal of Statistical Software*, *58*(6).

Meijer, R.R., Baneke, J. J. (2004). Analyzing psychopathology items: A case for nonparametric item response theory modeling. *Psychological Methods*, *9*(3), 354-368.

Mokken, R. J. (1971). *A theory and procedure of scale analysis*. De Gruyter, Berlin, Germany.

Nunnally, J. C. (1967). *Psychometric yheory*. New York: McGraw-Hill.

Panova, T., & Carbonell, X. (2018). Is smartphone addiction really an addiction? *Journal of Behavioral Addictions*, *7*(2), 252–259.

Pew Research Center (2019, August). Most U.S. teens who use cellphones do it to pass time, connect with others, learn new things. Retrieved from https://www.pewresearch.org/fact-tank/2019/08/23/most-u-s-teens-who-use-cellphones-do-it-to-pass-time-connect-with-others-learn-new-things/.

R: A language and environment for statistical computing (2013). R Core Team. R foundation for statistical computing, *1*, 409.

Rao, J. N. K., Scott, A. J. (1984). On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data. *Ann. Statist., 12*(1), 46-60.

Rizopoulos, R. (2006). Ltm: An r package for latent variable modeling and item response analysis. *Journal of Statistical Software*, *17*(1), 1–25.

Santor, D. A., & Ramsay, J. O. (1998). Progress in the technology of measurement: Applications of item response models. *Psychological Assessment, 10*(4), 345–359.

Stockdale, L. A., Coyne, S. M., & Padilla-Walker, L. M. (2018). Parent and child technoference and socioemotional behavioral outcomes: A nationally representative study of 10- to 20-year-old adolescents. *Computers in Human Behavior, 88,* 219–226.

Tan, C., M. Pamuk, & A. Donder (2013). *Loneliness and mobile phone*. *Procedia-Social and Behavioral Science, 103*, 606–611.

Tulane, S., Vaterlaus, J.M., & Beckert, T.E. (2018). A mixed methods examination of adolescents' reasons for pretending to text. *Journal of Adolescence, 69,* 175-179.

Twenge, J. M., Martin, G. N., & Campbell, W. K. (2018). Decreases in psychological well-being among American adolescents after 2012 and links to screen time during the rise of smartphone technology. *Emotion*, *18*(6), 765–780.

Walsh, S. P., White, K. M., & Young, R. M. D. (2010). Needing to connect: The effect of self and others on young people's involvement with their mobile phones. *Australian Journal of Psychology*, *62*(4), 194–203.

Zanon, C., Hutz, C., Yoo, H., & Hambleton, R. (2016). An application of item response theory to psychological test development. *Psicologia: Reflexão E Critica*, *29*(1).

Van der Ark, L. A. (2007). Mokken scale analysis in R. *Journal of Statistical Software, 20*(11), 1-19.