

Accepted Version of:

Gerosa T, Gui M and Büchi M (2021). Smartphone use and academic performance: a pervasiveness approach beyond addiction. *Social Science Computer Review*.
<https://doi.org/10.1177/08944393211018969>

Smartphone Use and Academic Performance: A Pervasiveness Approach Beyond Addiction

Tiziano Gerosa^{a*}, Marco Gui^a and Moritz Büchi^b

^a Department of Sociology and Social Research, University of Milano–Bicocca, Milano, Italy

^b Department of Communication and Media Research, University of Zurich, Zürich, Switzerland

* Corresponding author. E-mail: tiziano.gerosa@unimib.it

Abstract

Over the past decade smartphones have permeated all domains of adolescents' everyday lives, with research dominated by "smartphone addiction." This study compares one of the most used measures of smartphone addiction with a new alternative measure, the smartphone pervasiveness scale for adolescents (SPS-A), which focuses on the frequency of smartphone use at key social and physiological moments of daily life. A sample of 3,289 Italian high school students was used to validate the two constructs and compare their suitability for research on academic performance. SPS-A was moderately correlated with smartphone addiction, showed measurement invariance (across ethnic origins, parental education, and gender), and negatively predicted language and math test scores. SPS-A is a non-pathologizing instrument suitable to analyzing the role of smartphone use in academic achievement in combination with students' social background.

Keywords: smartphone pervasiveness, smartphone addiction, problematic smartphone use, academic performance, adolescents, well-being, scale validation, measurement invariance

From around 2010, smartphones acquired a predominant role as tools to access the internet among adolescents in developed countries (Mascheroni & Ólafsson, 2016). The increasing pervasiveness of such devices in young people's lives has raised public concerns about the ability of users to make a balanced use of them, alongside fears that smartphone use can become socially and psychologically harmful (see Turkle, 2016). Research has analyzed the association between smartphone usage and several outcomes (Elhai et al., 2016; Haug et al., 2015). However, the literature presents a number of theoretical and methodological issues, one of the most relevant being the way problematic smartphone use is conceptualized and, subsequently, measured (Amez & Baert, 2019). This study's main goal is to conceptually and statistically develop a measure of smartphone use that is sensitive to social background variables

and that does not pathologize adolescents' common behaviors yet can detect possible negative consequences.

The proposed measure of problematic smartphone use is based on the concept of "pervasiveness" and focuses on the frequency of use *at specific moments of the day*. These moments have been highlighted in the literature as particularly problematic for adolescents' social and physiological well-functioning (e.g. dinner with family, night hours, doing homework, etc.). The resulting measure, the Smartphone Pervasiveness Scale for adolescents (SPS-A), is administered and validated on an extensive sample of high school students ($N = 3,248$), also comparing its psychometric properties with those of the Smartphone Addiction Scale for adolescents in its short version (SAS-SV) (Kwon et al., 2013b).

The pertinence of the two scales as predictors of academic performance is then evaluated, controlling for a set of socio-demographic antecedents measured at both the individual and family level. Academic performance is a relevant outcome for adolescent's life chances, that is mostly measured through students' grade point average in the educational literature (see Amez & Baert, 2019). Although relatively easily obtainable by researchers, school grades suffer from subjectivity in the way teachers assign them and, in most studies, are self-reported by students, increasing measurement errors due to response biases. Administrative data overcomes these issues, offering standardized measures of the level of competences reached by students in national tests. We opted for standardized measures of learning performance in the two curricular disciplines Italian language and math, collected by the Italian Institute for educational evaluation (INVALSI) and merged with our survey data.

The main findings of the study are discussed in light of the need for social sciences to deal with new forms of social and educational inequalities in the mobile digital era, as well the reinforcement of existing ones.

Measuring Smartphone Use and Its Relationship with Academic Performance

Early studies on mobile technologies and education found a negative relationship between instant or text messaging and school performance, both at the actual and perceived level (Huang & Leung, 2009; Junco & Cotten, 2011). After 2010, when the massive diffusion of smartphones started in industrialized countries, research shifted its focus onto this new device, finding similar results (Lepp et al., 2014; Samaha & Hawi, 2016; Przybylski & Weinstein, 2013). In a recent literature review of the field, Amez and Baert (2019) confirm the negative relationship between smartphone use and learning outcomes. However, they also highlight a variety of ways in which smartphone use is operationalized, including self-reported measures of the quantity

of use, quantity of use from log data, and the addictive or problematic use self-reports. Such a variety of concepts and measurement approaches limits the comparability of the results, but also raises relevant theoretical questions about what exactly researchers need to measure when the relationship between smartphone use and academic performance is of interest.

Amount and frequency of Smartphone Use

Most of the previous literature on smartphone and academic performance exploits self-reported survey data about hours and/or minutes of use per day (e.g. Lepp et al., 2014; Chen & Ji, 2015). Although widely used, self-reports on the amount of time spent on this device have been recently criticized by cognitive psychologists, who found that time-perceptions can be significantly affected by prolonged use and rapid and consistently repeated checking behaviors (Wilcockson et al., 2018; Andrews et al., 2015). This is especially the case for open-ended questions focused on the exact time spent on smartphones and on the internet, often resulting in over- or under-estimations depending on a variety of respondent characteristics (Sewall et al., 2020; Ernala et al., 2020; Jurgens et al., 2019).

Using self-reported or observed measures of the overall amount of time spent by students using their devices also presents theoretical problems. On the one hand, it implies that it is the mere quantity of smartphone use that can make a difference, something which has been heavily criticized in recent studies on digital practices (Orben & Przybylski, 2019). On the other hand, it does not allow isolation of the specific moments and contexts of use that could most affect academic performance, preventing any substantial interpretation of the reason why smartphones have an ostensibly negative relationship with students' performance.

Research on the impact of digital uses has shown that the mere quantity of use may be inappropriate to predict negative educational outcomes (e.g. Liu, Baumeister, Yang, & Hu, 2019). Przybylski and Weinstein (2017) concluded that digital technology is not intrinsically harmful and may be advantageous in a connected world based on a large representative sample of British adolescents ($N = 120,115$). This has been confirmed by a following report based on PISA data showing that the relationship between digital media usage and learning achievements fits a hump-shaped curve with an optimum level (OECD, 2011). This is likely also since not all times of the day when the smartphone is used have the same capital enhancing or capital decreasing potential. For example, using this device frequently during class or late at night can negatively affect individuals' attentiveness, concentration and learning processes, while using it during free time spent alone may not have any negative consequence for school performance.

To overcome both the issue of individuals' estimation biases and of times of use, more recent studies have moved towards the collection of trace data through

online applications and platforms that monitor smartphone use (Felisoni & Godoi, 2018; Rosen et al., 2018). Kim and colleagues (2019), for example, used log data in a 14-week study with 84 first-year college students in Korea. They found an underestimation of in-class smartphone usage and a negative relationship with students' grades.

While promising, the collection of such data is challenging in several regards (e.g., Stier et al., 2020). First, obtaining and linking digital trace data to survey data can be complicated by the need to obtain personal information on participants together with their explicit and informed consent to use it for research purposes. Second, developing applications to access trace data from different smartphone operating systems is costly and does not grant access to equivalent information on all devices (Kreuter et al., 2020). An alternative way is to opt for data collection services already on the market. However, researchers' opportunities to obtain high quality data can be severely limited by their terms of use. These technological limitations, together with a potential lack of participant availability, can give rise to selection biases in the case of large, predefined samples (Al Baghal et al., 2019; Kreuter et al., 2020).

Problematic Smartphone Uses

A different strand of literature on smartphone use and academic performance focuses more specifically on problematic aspects (e.g. Harris et al., 2020). One of the most cited measures of problematic smartphone use for adolescents is the Smartphone Addiction Scale in its Short Version for adolescents (SAS-SV) (Kwon et al., 2013). The SAS-SV has been increasingly adopted in the literature on academic performance (Samaha & Hawi, 2016; Lee et al., 2015). However, smartphone addiction is not recognized as a psychiatric illness in the DSM-5 (American Psychiatric Association, 2013). The concept is measured by a series of self-report items that link smartphone overuse with daily-life disturbance and loss of control, relational difficulties, and physiological disorders. The SAS-SV has shown solid psychometric properties (Akin et al., 2014; Kwon et al., 2013) and was negatively related to youth's subjective well-being and educational outcomes (Lee et al., 2015; Samaha & Hawi, 2016). However, questionnaire items that frame smartphone usage as a disturbance introduce a number of methodological questions when its relationship with academic achievement is scrutinized.

First, it is not clear if a greater perception of problematic smartphone use is an indicator of actual misuse or of greater awareness of the smartphone's potential downsides. For example, a student who dedicates much time and effort to school activities might more readily report "having a hard time concentrating" when using the smartphone (Kwon et al., 2013). That is, an excellent student who devotes a lot of time to homework could have developed a high sensibility regarding smartphone

misuse. In this way, their perception of wasting time online could be more an indicator of anxiety about this behavior than of actual excessive smartphone use. This questions the use of similar scales as predictors of academic performance. Second, in so far as these scales focus on “addiction” like the SAS-SV, a problem emerges with the pathological implications of such a concept. One of the aims in the development of these measures was to offer diagnostic tools for clinical intervention. However, the focus of such scales on pathologies and addiction cannot discover negative effects on academic performance of less severe but presumably far more widespread problematic smartphone usage patterns. In this regard, some have suggested moving beyond the concept of addiction (Tokunaga, 2015; Panek, 2014), in favor of a more detailed analysis of the specific moments of adolescents’ daily life during which smartphone use could effectively become problematic, without necessarily being perceived as such (Cheever et al., 2018; Authors, 2020; Parent & Shapka, 2020). Surveys reveal that a majority of digital users experience problems in managing their online time and attention without simultaneously suffering from a pathological disorder (Ofcom, 2016; Rainie & Zickhur, 2015). Authors (2019) show empirically that digital overuse is not a pathological but mainly a social phenomenon. We consequently suggest investigating the relationship between smartphone use and academic achievement with measures of problematic use that rely on actual practices instead of feelings of pathological addiction.

Towards Smartphone Pervasiveness

In this paper we present an alternative instrument for measuring problematic smartphone use, the Smartphone Pervasiveness Scale for Adolescents (SPS-A), which aims to avoid both issues of individuals’ awareness and pathological framing. It consists of a scale measuring smartphone pervasiveness in relevant moments of adolescents’ daily life, an initial version of which has been already adopted in an extensive quantitative study (Authors, 2021; Authors, 2019). Addiction and pervasiveness differ in two main regards. Addiction concerns the inability to do without a smartphone despite the perception of negative effects on relevant daily-life activities, while pervasiveness is uniquely focused on objective frequency of use – in the sense that, although self-reported, it is free of judgement – during these relevant moments. Therefore, smartphone pervasiveness is simply a measure of frequency of use that *may* become problematic due to the moments of the day in which it is carried out. Such moments require the execution of demanding tasks and, more in general, concern offline activities that are generally not compatible with online activities (Authors, 2020).

From a social science perspective, this latter approach has the advantage of focusing on uses that could equally affect adolescents’ productivity and well-being

and, at the same time, be directly influenced by the social and cultural factors at the basis of digital inequality reproduction. Indeed, recent studies suggest that problematic use is the most recent domain of digital inequality to be significantly influenced by individuals' socio-cultural resources and family background (Authors, 2019; Authors, 2021). So far, the literature on problematic smartphone use has not disentangled the difference between the addiction- and pervasiveness-related measurement approaches. There is a need to deepen the analysis of their psychometric properties, their sensitivity to social and individual predictors and their potentially different impact on relevant outcomes such as academic performance. Researchers should be aware of how these two scales can be fruitfully employed depending on the specifics of their research question.

Smartphone pervasiveness focuses on daily life moments particularly relevant to adolescents' psychosocial well-being that could be affected by excessive smartphone use (e.g. dinner with family, talking with friends, homework, watching television or movies, and at night when awake). Media use at such moments could affect sleep quality and give rise to a loss of attentiveness and concentration, feelings of loneliness, social exclusion and depression. All symptoms which, in turn, could produce short- and long-term consequences on school performance and academic achievement (e.g. Dewald et al., 2010; Wolfson & Carskadon 2003; Mahapatra 2019; Wentzel, 2005; van der Schuur et al., 2015).

For example, social interaction with family and friends could be negatively influenced by frequent smartphone use, reducing empathy, closeness, and depth of conversation over time (Przybylski & Weinstein, 2013; Turkle, 2016). Regular attention to a smartphone while engaged in social interaction with a partner, friends, or even with parents, may encourage the same behavior to be reciprocated, leading to lower satisfaction with the relationship, feelings of loneliness and social exclusion (Authors, 2020; Chotpitayasunondh & Douglas, 2018; 2016; Roberts & David, 2016).

Turning to night-time, previous research has shown that prolonged exposure to smartphones can interfere with the sleep cycle (Christensen et al., 2016), while use at night was associated with sleep difficulties that can further the onset of depressive symptoms (Demirci et al., 2016; Lemola et al., 2015). Furthermore, smartphones can be used by adolescents while performing a wide range of other offline activities, increasing their propensity to multitask while studying or during leisure time (Hooft Graafland, 2018). Human information processing, however, is insufficient to attend to multiple stimuli and to perform simultaneous tasks (e.g. Wood et al., 2012), and research on multitasking has shown that media use while performing other activities is associated with increases in the time needed to carry out the primary task and loss of accuracy, attention and concentration (van der Schuur et al., 2015; Junco & Cotten, 2011; Kushlev et al., 2016).

In addition, recent studies have outlined new areas of interest in the analysis of smartphone daily life disturbance, in particular in morning routines and school time use. Empirical evidence highlights that a shorter time before first checking a smartphone in the morning is related to higher problematic use risks among youth (Haug et al., 2015) and negatively affects individual happiness and well-being (Hughes & Burke, 2018). At school, experimental studies have confirmed that smartphone use during lessons has a negative impact on academic performance (Felisoni & Godoi, 2018; Beland & Murphy, 2016). Indeed, smartphones represent a source of distraction that can interfere with activities such as note taking and undermine student learning (Waite et al., 2018).

Study Design and Aims

This study used an extensive sample of Italian adolescents to: (1) evaluate the psychometric properties of two measures of problematic smartphone use focused on pervasiveness (SPS-A) and addiction (SAS-SV), respectively; and (2) compare their suitability and sensitivity for research on academic performance. The first aim concerns problematic smartphone use *measurement issues*, while the second deals with its *relationship with student academic performance*, accounting for potential confounders at individual and family levels (Figure 1).

Regarding the first goal, the *dimensionality* of SPS-A and SAS-SV was explored without imposing any constraints on the number of factors or the specific items they should comprise to test whether SPS-A and SAS-SV are interchangeable or distinct – even if interrelated – measures of problematic smartphone use. We evaluated their *factorial validity* in a pooled confirmatory framework and then tested whether the observed items effectively covered the best fitting model specification from previous analysis. That is, we postulated specific relations between the observed items and their underlying factors to evaluate how well the empirical data fit more restrictive and model specifications defined *a priori*. We investigated the *measurement invariance* (MI) of the two constructs across students grouped by relevant predictors of academic performance, such as gender, ethnic origin, and educational background (e.g. Marks, 2008; Davis-Kean, 2005; Glick & Hohmann-Marriot, 2007). Through MI we assessed the equivalence of the equations used to create pervasiveness and addiction factor scores across the sub-populations of interest (Meredith, 1993; Vandenberg & Lance, 2000).

[Figure 1]

The second stage of the study concerns the relationship between the two measures and students' academic performance. We first investigated if known predictors of school performance (gender, parents' educational level, ethnic origins) are also related to measures of problematic smartphone use by analyzing *cross-group*

latent mean differences. This allowed us to evaluate potential confounding in the studied relationship. For instance, we know from previous studies that females achieve lower average scores in math tests compared with their male classmates and, at the same time, girls are more exposed than boys to smartphone-related problems (e.g. Kwon et al., 2013). Therefore, a predominance of females in the sample without controlling for it would lead to overestimating the direct and negative association between problematic smartphone use and math test scores. In the final step of the analysis, we estimated the *predictive power* of SPS-A and SAS-SV on students' performance in Italian language and math standardized tests, controlling for the potential confounders identified in the previous analysis.

Method

Participants

The analyses were carried out on data from the second wave of data collection of the [anonymized] project (Authors, 2020). This project involved an extensive survey on all 10th grade students in 18 high schools located in two neighboring districts of the [anonymized] region, northern Italy. In May 2018, students were invited to fill out an online questionnaire. The survey was administered in the school's computer labs during class time and guided by external observers. At the end of the data collection process, 3,289 students of the 3,659 officially registered students completed the survey, reaching a final response rate of 90%.¹ We also obtained an additional set of administrative information on students' school performance by merging our data with those collected by the National Institute for the Evaluation of the Education System (INVALSI). INVALSI represents the reference institution in Italy for the assessment of students' achievement since 2007. During the INVALSI year-end survey activities, students are requested to perform computer-based tests aimed at measuring their competence in the two curricular disciplines of Italian language and math. Participation in the survey is compulsory for the entire population of students at various levels of education, including 10th grade. We first asked schools to inform parents of the initiative and then to sign an agreement that allowed us to access data collected by INVALSI, in accordance with data protection rules. We then obtained students' anonymous identification codes that were sent to the INVALSI statistical office, who in turn sent us the standardized test scores for both disciplines. Finally, we linked INVALSI standardized test scores to our data, reaching a coverage of 80% of the population under study. Detailed sample characteristics are reported in Table 1.

¹ The missing respondents were mainly those absent on the dates of the survey and any rescheduled dates, together with those being withdrawn from schools or giving up their course during the year.

[Table 1]

Measures

The project questionnaire covered students' socio-demographic characteristics, digital competence, attitudes toward digital technologies, daily smartphone usage habits, and problematic smartphone usage (Authors, 2020). For this last topic, the questionnaire included both the SPS-A and SAS-SV measure.

Smartphone Pervasiveness Scale for adolescents (SPS-A). The SPS-A consists of a set of items asking students how frequently they use their personal device in different moments of the day on a 4-point ordinal scale ranging from 1 (*never*) to 4 (*always*). In its original version (Authors, 2021; Authors, 2019), the SPS-A was focused on five daily life moments particularly relevant to adolescents' psychosocial well-being that, according to the literature, could be affected by excessive smartphone use: dinner with family, time with friends, homework, watching television or movies, and at night when awake. This 5-item scale was first administered and validated on a sample of 4,675 Italian upper secondary school students, obtaining satisfactory results both in terms of factorial validity and internal consistency. More recently, a revised version of SPS-A was developed based on the empirical evidence on the consequences of consistently checking the smartphone close to sleep time and during lessons at school (Authors, 2020). Two further items were added to the original scale: a first item asks respondents how often they use their smartphone first thing in the morning when they wake up, while the second focuses on frequency of use at school during class. This revised version of the SPS-A was validated following a two-stage approach of item and scale development. The items were first selected and refined according to their content validity. The resulting 7-item scale was administered to a sample of 3,361 upper secondary students and proved to be unidimensional, valid and sufficiently reliable. A detailed description of the item selection and the scale validation procedures and results are reported in the supplementary material.

Smartphone Addiction Scale – short version for adolescents (SAS-SV). The SAS-SV focuses on respondents' perceptions of excessive use and the degree of discomfort arising from it (Kwon et al., 2013b). SAS-SV is measured with 10 items on a 6-point scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). The items were originally reflective of different subscales (see supplementary materials): daily life disturbance (DD), withdrawal (WI), overuse (OV), tolerance (TO) and disposition to cyberspace-oriented relationships (CR). The English and Italian SAS-SV are reported in the supplementary material.

Students' standardized test scores in Italian language (TSI) and math (TSM). TSI and TSM were extracted from a computer-based survey carried out in May 2018 on all 10th grade students. Students' scores at the INVALSI TSI and TSM are highly

reliable and directly comparable across the entire population, overcoming the limit of subjectivity in the evaluation of students' competences and the risk of measurement errors due to self-reported grade point averages. Students' scores in both tests are standardized at the national level to a mean of 200 and standard deviation of 50. In our sample, students' scores are less dispersed and slightly above the national average in both disciplines (Table 1).

Covariates. Participants were asked to report a set of socio-demographic attributes: gender, ethnic origin, and the education level of their parents. Students' gender and ethnic origin were recorded as dichotomous variables, distinguishing males from females and natives from the first and second generations of immigrants, respectively. The highest level of education among parents (dominance criterion) was queried and then recoded into low-educated (up to middle school diploma), middle-educated (high school diploma) and highly educated (bachelor's degree or higher).

Data analysis

Due to the specificity of our research interests, all the analyses presented in the study were conducted on the subsample of participants who reported owning a smartphone ($n = 3,251$, around 99% of the overall sample). A preliminary evaluation of the degree of multivariate normality of all problematic smartphone use items was performed through the Mardia test. Results demonstrated significant skewness and kurtosis, thus estimation methods specifically designed for non-normally distributed ordered categorical data were used. Exploratory and single-group confirmatory factor analysis (EFA and CFA) models aimed at evaluating the dimensionality and the factorial validity of SAS-SV and SPS-A were estimated with weighted least squares (WLS) and mean- and variance-adjusted chi-square test statistic (χ^2). Model goodness of fit was considered to adequately explain the empirical data when the χ^2 probability was higher than 0.05. However, previous research showed that χ^2 values could be inflated with large sample sizes, leading to erroneous conclusions about the model fit to the empirical data and the tightness of the conclusions reached toward the cross-group comparisons (Chen, 2007). The following set of alternative fit indices was then included: the root mean squared error of approximation (RMSEA); the comparative fit index (CFI); the Tucker–Lewis index (TLI).² Multi-group CFA models for measurement invariance testing were estimated with a robust maximum

² According to the rule of thumb, values of RMSEA lower than 0.05 indicate a good fit, while values lower than 0.08 are acceptable (Hu & Bentler, 1999). CFI and TLI higher than 0.90 and 0.95 represent acceptable and good fit, respectively.

likelihood (ML) estimator, looking at $\Delta\chi^2$ and variations in McDonald's non-centrality index (Mc) and CFI to compare nested models (see Sass et al., 2014).³

To avoid the problems associated with testing EFA and CFA on the same data set (Kline 2015), we divided our sample using a random half-splitting procedure, and conducted dimensionality and validity testing on two independent subsamples (Table 1). The number of potential factors underlying the list of selected items and the ways in which they relate were investigated on Subsample 1 applying multiple criteria for the evaluation of construct dimensionality. We first focused on the screen plot with eigenvalues, adopting the elbow rule and a cutoff of 1 to identify the number of factors to retain (Stevens, 2012). We then carried out incremental (EFA) models with Geomin rotation by progressively increasing the estimated factors until reaching the number suggested by previous analysis. Once the best fitting solution was identified, we examined factor loadings and inter-factor correlations in terms of substantive meaning.

To evaluate the factorial validity, we estimated a series of alternative CFA models on Subsample 2. In the baseline model, SPS-A and SAS-SV items loaded only on their respective factors and no item residuals were allowed to covary. We started from the global fit of the model and used modification indices to judge the improvements if previously omitted correlated residuals, i.e., constrained to zero, are freely estimated. To identify which of the estimated modification indices should be considered as a potentially relevant residual covariance, we opted for a threshold of 0.2, meaning that at least 4% of the item variance not explained by the SAS-SV latent factor is shared with other items.

Measurement invariance was assessed on the overall sample through a series of hierarchical nested models imposing an increasing number of parameter constraints across groups of interest (multi-group confirmatory factor analysis; MGCFAs). Without sufficient MI, differences detected are at least partially due to measurement biases rather than substantive differences (Büchi, 2016). Our analytical strategy adapted the seminal work of Widaman and Reise (1997) to comparing model parameters across groups distinguished by gender, ethnic origin, and parent educational level. To achieve this goal and proceed with latent means comparisons, at least partial scalar invariance between items intercepts must be established (Kline, 2015), i.e., at least one additional intercept other than that of the reference item needs to be equivalent across groups. We identified baseline models for each subgroup of respondents and tested configural, metric, residual covariances, and

³ A probability higher of 0.05 for $\Delta\chi^2$ p-value means that the more restricted model fits the data almost as equally as the less restricted model. Values of ΔMc and ΔCFI less than or equal to 0.02 and 0.01 are considered sufficient support of cross-group equivalence (Cheung & Rensvold, 2002).

scalar invariance across them.⁴ Configural invariance evaluates whether the basic organization of the constructs, namely item loadings on each latent factor, applies to all groups. Metric invariance focuses on the equivalence of factor loadings, while residual covariances invariance is used to assess whether common residual covariances in the baseline models operate equivalently. Finally, scalar invariance was tested by constraining the item intercepts to be equivalent across the groups of interest. To then compare latent mean scores across groups we constrained one of them to 0 (reference group) and left all the others free to vary (Kline, 2015). Following this procedure, the factor means of the freely estimated groups summarize the differences between the reference and other groups' latent factor means. In other words, all the parameters left free to vary represent the true difference in latent means as compared to the reference group.

Finally, we simultaneously regressed TSI and TSM on SPS-A and SAS-SV using data from the overall sample of smartphone owners to investigate the relationships between the two constructs of problematic use and students' school performance. Our model specification included potential socio-demographic confounders previously examined, introducing them as covariates of SPS-A and SAS-SV. More technical details on the methodological choices underlying fit evaluation, rotation criteria selection, factor validity and invariance testing are given in the supplementary material.

Results

Stage 1: Dimensionality

The scree plot (supplementary material) as a first indication of dimensionality suggested retaining more factors than the two initially hypothesized. The first three factors featured eigenvalues higher than 1 and the graph's elbow was positioned exactly in correspondence with the third. Similar evidence was obtained from the alternative fit statistics of incremental EFA models. The one-factor solution ($\chi^2 = 3037$; $df = 119$; χ^2 p-val = 0.000; RMSEA = 0.123 [0.120-0.127]; CFI = 0.812; TLI = 0.785), as well as the two-factor solution ($\chi^2 = 1440$; $df = 103$; χ^2 p-val = 0.000; RMSEA = 0.090 [0.086-0.094]; CFI = 0.914; TLI = 0.886) did not fit the data well and were rejected. Conversely, the three-factor solution produced RMSEA, CFA and TLI values indicating a more than acceptable fit to the data ($\chi^2 = 739$; $df = 88$; χ^2 p-val = 0.000; RMSEA = 0.064 (0.063-0.072); CFI = 0.958; TLI = 0.935).

⁴ Although required for full factorial invariance (Meredith, 1993), we avoided the last step of the Widaman and Reise strategy based on the equivalence of item residuals (strict invariance). This choice depends on the fact that residuals are not part of the latent factor and therefore becomes inconsequential for the interpretation of latent mean differences across groups (Vandenberg & Lance, 2000).

[Table 2]

Looking at the loadings of the three-factor solution (Table 2), the two separate latent dimensions of problematic smartphone use underlying the SPS-A and SAS-SV items were evident. Indeed, for each of the two constructs the parameter estimates exhibited values above the 0.5 threshold and no signs of overlap were detected. A third factor of less obvious interpretation was instead defined by the cross-loadings of items SAS2, SAS3, SAS5 and SAS8, respectively, focused on smartphone overuse and its negative effects on productivity (see the list of items in the supplementary material). The specificity of these items compared to the rest of the SAS-SV battery suggests that this third factor could simply summarize individuals' (lack of) sensitivity toward smartphone daily life disturbance. In line with this hypothesis, we found a moderate positive inter-factor correlation between SPS-A and SAS-SV, while, conversely, the third factor resulted in a null correlation with each of the two measures of problematic use.

Stage 2: Factorial Validity

Results of the analysis on the baseline model did not yield a satisfactory goodness of fit ($\chi^2 = 1316$; $df = 118$; χ^2 p-val = 0.000; RMSEA = 0.079 [0.076–0.083], CFI = 0.920, TLI = 0.908). Based on the model modification strategy, we identified four misspecified residual covariances (SAS2-SAS3, SAS2-SAS8, SAS3-SAS8, and SAS5-SAS8), ranging from 0.264 to 0.483 and dealing – as expected – with smartphone overuse and its negative effects on productivity. The hypothesized model was then gradually respecified to test whether the addition of each of these parameters corresponded to a significant improvement in the global fit indices. The final model specification resulted in values of RMSEA, CFI and TLI indicating a close fit to the data ($\chi^2 = 649$; $df = 114$; χ^2 p-val = 0.000; RMSEA = 0.054 [0.050–0.058], CFI = 0.96, TLI = 0.957). Some SAS-SV items – those related to smartphone overuse and productivity disturbances – are not adequately explained by the construct of addiction and require the introduction of correlated error parameters.

Stage 3: Measurement Invariance

In short, with for example $\Delta CFI < .01$ and $\Delta Mc < .02$, the results of MGCFA confirmed that SPS-A and SAS-SV represent two fully invariant constructs across ethnic origins and parental education, satisfying all the necessary prerequisites for analysis of the latent factor means (Table 3). SPS-A was also invariant for gender, whereas SAS-SV proved to be only partially invariant.

[Table 3]

Table 3 shows that configural invariance, as well as metric and residual covariance invariance, were all confirmed. Despite the marked inflation produced on χ^2 by our large sample size, RMSEA, CFI, TLI and McNCI values consistently

indicated at least an acceptable fit, while values of ΔCFI and $\Delta McNCI$ remained below the recommended threshold in all of the first three model comparisons. Conversely, scalar invariance testing produced a substantial decrease in model fit ($\Delta\chi^2$ p -val $< 0,05$; $\Delta CFI = - 0.016$; $\Delta McNCI = - 0.025$), with modification indices suggesting that items SAS8 and SAS9 were non-invariant across gender. After releasing the equality constraints on these two non-invariant parameters (Model 4.1), the magnitude of alternative fit indices variations reduced enough to support partial scalar invariance for gender ($\Delta CFI = - 0.009$; $\Delta McNCI = - 0.013$).

Stage 4: Latent Means Comparisons by Socio-Demographic Characteristics

Based on the stage 3 results, we compared SPS-A and SAS-SV factor means for students' gender, ethnic origins, and parental education background. Table 4 show that SPS-A was significantly predicted by parental education and ethnic origins, but not by gender. On the other hand, SAS-SV was not significantly associated with family background, while female students on average scored higher on SAS-SV than males. To evaluate whether SAS-SV latent factor means are impacted by non-invariant items, we carried out a sensitivity analysis comparing the parameter estimates of the gender multi-group fully invariant model with those of the partially invariant model. Results of this additional analysis underline that the higher propensity of female students to declare smartphone overuse and dependency contributes to an inflation of the estimates: in the fully invariant model we see a 29% increase in the latent means difference between males and females compared with the partially invariant model.

[Table 4]

Stage 5: Academic Performance

A latent model in which TSI and TSM were regressed on both SPS-A and SAS-SV was finally carried on the overall sample to investigate the relationships between the two constructs of problematic smartphone use and students' school performance. Our model specification included potential socio-demographic confounders previously examined, introducing them as covariates of SPS-A and SAS-SV. This resulted in an acceptable fit ($\chi^2 = 1308.2$; $df = 212$; $p = 0.000$; $RMSEA = 0.043[0.031-0.045]$; $CFI = 0.914$; $TLI = 0.900$). We found that levels of SPS-A significantly and negatively predicted test scores in both disciplines, indicating that the less students report their smartphone as pervasive, the better they did at school (Table 5). We found lower and non-significant associations between SAS-SV and students' performance in Italian language and math.

[Table 5]

Discussion

Dimensionality testing with EFA supported the hypothesis that SPS-A and SAS-SV are separate but related measures of problematic smartphone use. The third factor in the analysis might merely incorporate residual covariances across the SAS-SV items more specifically focused on perceived smartphone overuse and its negative effects on productivity rather than addiction and withdrawal issues. These items have in common that they elicit perceptions of habits (and their consequences) that addicted respondents may refuse to accept as their own or simply do not want to report, while respondents with non-problematic use may have a higher sensitivity and anxiety toward the consequences of smartphone use on daily life activities. The specificity of these items compared with the rest of the SAS-SV suggests that this third factor could simply summarize their shared variance unaccounted for by the constructs of problematic smartphone use. Supporting this hypothesis, the two-factor model with correlated scales and residual covariances was the best fitting solution.

In this model, SPS-A and SAS-SV represent distinct but not orthogonal constructs. Thus, addiction and pervasiveness are empirically overlapping, and yet they are not interchangeable measures of problematic smartphone use. Moreover, the patterns of residual variance show the advantage of being able to account for measurement error in a latent modeling framework and should discourage researchers from using a simple mean score of SAS-SV items. Our findings of partial measurement invariance mean that SPS-A and SAS-SV latent mean scores comparisons by gender may be biased, particularly because the decision to relax the constraints for non-invariant items could lead to comparability problems between factor scores (e.g., Sass, 2011). Sensitivity analysis in Stage 4 revealed that the magnitude of gender differences in the smartphone addiction scores must be interpreted with caution.

Finally, stage 5 results suggest reconsidering previous evidence on the negative relationship between SAS-SV and school performance: smartphone addiction perceived by students did not significantly predict their school performance once controlled for smartphone pervasiveness at specific moments of the day and socio-demographic covariates. SAS-SV was not associated with social and ethnic origins of the students, while SPS-A showed significant differences for both. SAS-SV yielded gender differences as reported in previous research (Kwon et al., 2013), while SPS-A was not significantly affected by gender. Aside from the higher sensitivity of females toward the issue of problematic smartphone use and its daily life disturbance, the gender gap frequently reported in the problematic smartphone use literature is more generally attributable to differences in males and females' motivations and types of use, rather than frequency of use at specific

moments of the day (*Authors, 2020; Chen et al., 2017*). Analyzing the robustness of the relationships between the two scales and students' performance in Italian language and math showed that only smartphone pervasiveness was negatively associated with students' standardized test scores in both disciplines.

Conclusion

This article introduced an alternative conceptualization and measurement of problematic smartphone use, validated and applied in a sample of 3,289 10th grade students in Italy. The SPS-A represents a new resource for future social science research on the impact of smartphones on school performance. It is not a measure of pathological smartphone usage behaviors, nor it is a measure of (subjective) overuse; instead, it focuses on the frequency of use at specific key moments of the day. Its crucial features are factorial validity – meaning that smartphone pervasiveness is a single latent dimension adequately reflected in the frequency of smartphone use in a set of daily situations – and measurement invariance across relevant groups. Thus, the properties of the SPS-A enable meaningful comparisons between socio-demographic groups of interest for research on the role of smartphone use for educational attainment and inequalities among adolescents.

Given our results, SPS-A presents a more suitable choice for researchers interested in studying the relationship between smartphone usage and school performance, even if controlling for other factors associated with the reproduction of educational inequalities. Moreover, since SPS-A is associated with socio-demographic variables of family background, it may represent a more theoretically meaningful choice for communication researchers investigating how family habits and cultural norms regarding digital media are transmitted unequally. SPS-A allows researchers interested in problematic smartphone use and its outcomes to go beyond the prevalent “addiction” perspective. As the integration of smartphones into everyday communication is a societal and not only a psychological process, researchers may want to move outside of a pathology-oriented framework. SPS-A better shows the social drivers of problematic smartphone use that does not derive from individual psychological problems but from socialization through family and peer-learned communication habits.

Limitations and further research

Some limitations of the study need to be acknowledged. First, it SPS-A still relies on adolescents' self reports, which may be subject to estimation biases. Self-report measures of smartphone use are indeed significantly affected by the difficulty to

report frequent and habitual checking behaviors (Wilcockson et al., 2018; Andrews et al., 2015).

Likert scale measures such as the SPS-A could be considered less sensitive to this issue compared to open-ended questions on the amount of time spent on the device (Sewall et al., 2020; Ernala et al., 2020). However, research in this field is still scarce and more empirical evidence is needed. Future studies should then invest more in trace data assessments and incorporate objective measures of frequency of smartphone use in specific moments of the day – ideally in combination with relevant (self-reported) background and outcome variables – to strengthen the evaluation of SPS-A criterion validity.

Second, the robustness and direction of the relationship we found across SPS-A and school performance remains somewhat speculative. Even though the analysis we presented accounted for some of the most relevant predictors of smartphone use and school performance at the individual and family level, potentially omitted confounders and reverse causality deriving from the observational nature of our data remain (Rohrer, 2018). Although it is generally assumed that problematic smartphone use leads to worse school performance, an increase in smartphone use may also be the result of a decline in school performance. Recent advancements in longitudinal research address these issues by modeling the auto-regressive effects of predictors and outcomes through random-effect techniques that control for both intra- and inter-individual changes over time (e.g. Selig & Little, 2012). Scholars interested in deepening the causal analysis of smartphone pervasiveness and academic performance are thus encouraged to direct their efforts toward the application of such techniques.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the *[anonymized]*.

Ethical consent

The study was conducted in compliance with the regulations issued by the Ethics Committee of the University of *[anonymized]* (protocol number 372).

References

- Al Baghal, T., Sloan, L., Jessop, C., Williams, M. L., & Burnap, P. (2020). Linking Twitter and survey data: The impact of survey mode and demographics on consent rates across three UK studies. *Social Science Computer Review, 38*(5), 517-532.
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: Author.
- Amez, S., & Baert, S. (2019). *Smartphone Use and Academic Performance: A Literature Review* (SSRN Scholarly Paper ID 3483961). Social Science Research Network. <https://papers.ssrn.com/abstract=3483961>
- Andrews, S., Ellis, D. A., Shaw, H., & Piwek, L. (2015). Beyond self-report: tools to compare estimated and real-world smartphone use. *PLOS one, 10*(10), e0139004.
- Beland, L. P., & Murphy, R. (2016). Ill communication: technology, distraction & student performance. *Labour Economics, 41*, 61-76.
- Büchi, M. (2016). Measurement invariance in comparative Internet use research. *Studies in Communication Sciences, 16*(1), 61-69. <https://doi.org/10.1016/j.scoms.2016.03.003>
- Cheever, N. A., Moreno, M. A., & Rosen, L. D. (2018). When Does Internet and Smartphone Use Become a Problem? In M. A. Moreno & A. Radovic (Eds.), *Technology and Adolescent Mental Health* (pp. 121-131). Springer International Publishing. https://doi.org/10.1007/978-3-319-69638-6_10
- Chen, C., Zhang, K. Z., Gong, X., Zhao, S. J., Lee, M. K., & Liang, L. (2017). Examining the effects of motives and gender differences on smartphone addiction. *Computers in Human Behavior, 75*, 891-902.
- Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464-504. <https://doi.org/10.1080/10705510701301834>
- Chen, R. S., & Ji, C. H. (2015). Investigating the relationship between thinking style and personal electronic device use and its implications for academic performance. *Computers in Human Behavior, 52*, 177-183.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5
- Chotpitayasunondh, V., & Douglas, K. M. (2018). The effects of “phubbing” on social interaction. *Journal of Applied Social Psychology, 48*(6), 304-316.
- Chotpitayasunondh, V., & Douglas, K. M. (2016). How “phubbing” becomes the norm: The antecedents and consequences of snubbing via smartphone. *Computers in Human Behavior, 63*, 9-18.
- Christensen, M. A., Bettencourt, L., Kaye, L., Moturu, S. T., Nguyen, K. T., Olgin, J. E., ... & Marcus, G. M. (2016). Direct measurements of smartphone screen-time: relationships with demographics and sleep. *PLOS ONE, 11*(11).
- Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of family psychology, 19*(2), 294.

- Demirci, K., Akgönül, M., & Akpınar, A. (2015). Relationship of smartphone use severity with sleep quality, depression, and anxiety in university students. *Journal of Behavioral Addictions, 4*(2), 85-92.
- Dewald, J. F., Meijer, A. M., Oort, F. J., Kerkhof, G. A., & Bögels, S. M. (2010). The influence of sleep quality, sleep duration and sleepiness on school performance in children and adolescents: A meta-analytic review. *Sleep Medicine Reviews, 14*(3), 179-189.
- Elhai, J. D., Levine, J. C., Dvorak, R. D., & Hall, B. J. (2016). Fear of missing out, need for touch, anxiety and depression are related to problematic smartphone use. *Computers in Human Behavior, 63*, 509-516.
- Felisoni, D. D., & Godoi, A. S. (2018). Cell phone usage and academic performance: An experiment. *Computers & Education, 117*, 175-187.
- Glick, J. E., & Hohmann-Marriott, B. (2007). Academic performance of young children in immigrant families: The significance of race, ethnicity, and national origins. *International Migration Review, 41*(2), 371-402.
- Hampton, K. N. (2019). Social Media and Change in Psychological Distress Over Time: The Role of Social Causation. *Journal of Computer-Mediated Communication, 24*(5), 205-222. <https://doi.org/10.1093/jcmc/zmz010>
- Harris, B., Regan, T., Schueler, J., & Fields, S. A. (2020). Problematic mobile phone and smartphone use scales: A systematic review. *Frontiers in Psychology, 11*.
- 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.
- Hooft Graafland, J. (2018). New technologies and 21st century children: recent trends and outcomes. OECD Education Working Papers, 179. <https://doi.org/10.1787/e071a505-en>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Huang, H., & Leung, L. (2009). Instant messaging addiction among teenagers in China: Shyness, alienation, and academic performance decrement. *CyberPsychology & Behavior, 12*(6), 675-679. <https://doi.org/10.1089/cpb.2009.0060>
- Hughes, N., & Burke, J. (2018). Sleeping with the frenemy: How restricting 'bedroom use' of smartphones impacts happiness and wellbeing. *Computers in Human Behavior, 85*, 236-244.
- Junco, R., & Cotten, S. R. (2011). Perceived academic effects of instant messaging use. *Computers & Education, 56*(2), 370-378. <https://doi.org/10.1016/j.compedu.2010.08.020>
- Jürgens, P., Stark, B., & Magin, M. (2020). Two half-truths make a whole? On bias in self-reports and tracking data. *Social Science Computer Review, 38*(5), 600-615.
- Kim, I., Kim, R., Kim, H., Kim, D., Han, K., Lee, P. H., ... & Lee, U. (2019). Understanding smartphone usage in college classrooms: A long-term measurement study. *Computers & Education, 141*, 103611.
- Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling, Fourth Edition*. Guilford Publications.

- Kreuter, F., Haas, G. C., Keusch, F., Bähr, S., & Trappmann, M. (2020). Collecting survey and smartphone sensor data with an app: Opportunities and challenges around privacy and informed consent. *Social Science Computer Review*, *38*(5), 533-549.
- Kushlev, K., Proulx, J., & Dunn, E. W. (2016, May). " Silence Your Phones" Smartphone Notifications Increase Inattention and Hyperactivity Symptoms. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 1011-1020).
- Kwon, M., Kim, D. J., Cho, H., & Yang, S. (2013b). The smartphone addiction scale: development and validation of a short version for adolescents. *PLOS ONE*, *8*(12).
- Kwon, M., Lee, J.-Y., Won, W.-Y., Park, J.-W., Min, J.-A., Hahn, C., Gu, X., Choi, J.-H., & Kim, D.-J. (2013). Development and Validation of a Smartphone Addiction Scale (SAS). *PLOS ONE*, *8*(2), e56936. <https://doi.org/10.1371/journal.pone.0056936>
- Lee, J., Cho, B., Kim, Y., & Noh, J. (2015). Smartphone Addiction in University Students and Its Implication for Learning. In G. Chen, V. Kumar, Kinshuk, R. Huang, & S. C. Kong (Eds.), *Emerging Issues in Smart Learning* (pp. 297-305). Springer. https://link.springer.com/chapter/10.1007/978-3-662-44188-6_40
- Lemola, S., Perkinson-Gloor, N., Brand, S., Dewald-Kaufmann, J. F., & Grob, A. (2015). Adolescents' electronic media use at night, sleep disturbance, and depressive symptoms in the smartphone age. *Journal of Youth and Adolescence*, *44*(2), 405-418.
- Lepp, A., Barkley, J. E., & Karpinski, A. C. (2014). The relationship between cell phone use, academic performance, anxiety, and satisfaction with life in college students. *Computers in Human Behavior*, *31*, 343-350.
- Liu, D., Baumeister, R. F., Yang, C., & Hu, B. (2019). Digital Communication Media Use and Psychological Well-Being: A Meta-Analysis. *Journal of Computer-Mediated Communication*, *24*(5), 259-273. <https://doi.org/10.1093/jcmc/zmz013>
- Mahapatra, S. (2019). Smartphone addiction and associated consequences: Role of loneliness and self-regulation. *Behaviour & Information Technology*, *38*(8), 833-844.
- Mascheroni, G., & Ólafsson, K. (2016). The mobile Internet: Access, use, opportunities and divides among European children. *New Media & Society*, *18*(8), 1657-1679.
- Marks, G. N. (2008). Accounting for the gender gaps in student performance in reading and mathematics: evidence from 31 countries. *Oxford Review of Education*, *34*(1), 89-109.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525-543. <https://doi.org/10.1007/BF02294825>
- Orben, A., & Przybylski, A. K. (2019). Screens, teens, and psychological well-being: Evidence from three time-use-diary studies. *Psychological science*, *30*(5), 682-696.
- Panek, E. (2014). Left to their own devices: College students' "guilty pleasure" media use and time management. *Communication Research*, *41*(4), 561-577.
- Parent, N., & Shapka, J. (2020). Moving beyond addiction: An attachment theory framework for understanding young adults' relationships with their smartphones. *Human Behavior and Emerging Technologies*, *2*(2), 179-185. <https://doi.org/10.1002/hbe2.180>
- Przybylski, A. K., & Weinstein, N. (2013). Can you connect with me now? How the presence of mobile communication technology influences face-to-face conversation quality. *Journal of Social and Personal Relationships*, *30*(3), 237-246. <https://doi.org/10.1177/0265407512453827>

- Przybylski, A. K., & Weinstein, N. (2017). A large-scale test of the goldilocks hypothesis: quantifying the relations between digital-screen use and the mental well-being of adolescents. *Psychological Science, 28*(2), 204-215.
- Rainie, L., & Zickuhr, K. (2015). Americans' views on mobile etiquette. *Pew Research Center*. <http://www.pewinternet.org/2015/08/26/americans-views-on-mobile-etiquette>
- Roberts, J. A., & David, M. E. (2016). My life has become a major distraction from my cell phone: Partner phubbing and relationship satisfaction among romantic partners. *Computers in Human Behavior, 54*, 134-141.
- Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science, 24*(2), 251524591774562. <https://doi.org/10.1177/2515245917745629>
- Samaha, M., & Hawi, N. S. (2016). Relationships among smartphone addiction, stress, academic performance, and satisfaction with life. *Computers in Human Behavior, 57*, 321-325. <https://doi.org/10.1016/j.chb.2015.12.045>
- Sass, D. A. (2011). Testing Measurement Invariance and Comparing Latent Factor Means Within a Confirmatory Factor Analysis Framework. *Journal of Psychoeducational Assessment, 29*(4), 347-363. <https://doi.org/10.1177/0734282911406661>
- Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating Model Fit With Ordered Categorical Data Within a Measurement Invariance Framework: A Comparison of Estimators. *Structural Equation Modeling: A Multidisciplinary Journal, 21*(2), 167-180. <https://doi.org/10.1080/10705511.2014.882658>
- Selig, J.P. & Little, T.D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In: Laursen B, Little TD and Card NA (eds) *Handbook of Developmental Research Methods*. New York: The Guilford Press, pp. 265-278.
- Sewall, C. J., Bear, T. M., Merranko, J., & Rosen, D. (2020). How psychosocial well-being and usage amount predict inaccuracies in retrospective estimates of digital technology use. *Mobile Media & Communication, 2050157920902830*.
- Stevens, J. P. (2012). *Applied Multivariate Statistics for the Social Sciences*. Routledge.
- Stier, S., Breuer, J., Siegers, P., & Thorson, K. (2020). Integrating survey data and digital trace data: key issues in developing an emerging field. *Social Science Computer Review, 38*(5), 503-516.
- Tokunaga, R. S. (2015). Perspectives on Internet addiction, problematic Internet use, and deficient self-regulation: Contributions of communication research. *Annals of the International Communication Association, 39*(1), 131-161.
- Turkle, S. (2016). *Reclaiming Conversation: The Power of Talk in a Digital Age*. Penguin Press.
- Vandenberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods, 3*(1), 4-70. <https://doi.org/10.1177/109442810031002>
- Van Der Schuur, W. A., Baumgartner, S. E., Sumter, S. R., & Valkenburg, P. M. (2015). The consequences of media multitasking for youth: A review. *Computers in Human Behavior, 53*, 204-215.

- Waite, B. M., Lindberg, R., Ernst, B., Bowman, L. L., & Levine, L. E. (2018). Off-task multitasking, note-taking and lower-and higher-order classroom learning. *Computers & Education, 120*, 98-111.
- Wentzel, K. R. (2005). Peer relationships, motivation, and academic performance at school. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of Competence and Motivation*, 279–296. New York: Guilford Press.
- Widaman, K. F. & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: applications in the substance use domain. In K. J. Bryant & M. Windle (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281-324). Washington, DC: American Psychological Association.
- Wilcockson, T. D., Ellis, D. A., & Shaw, H. (2018). Determining typical smartphone usage: What data do we need? *Cyberpsychology, Behavior, and Social Networking, 21*(6), 395-398.
- Wolfson, A. R., & Carskadon, W. (2003). Understanding adolescents' sleep patterns and school performance. *Sleep Med. Rev, 7*, 491-506.
- Wood, E., Zivcakova, L., Gentile, P., Archer, K., De Pasquale, D., & Nosko, A. (2012). Examining the impact of off-task multi-tasking with technology on real-time classroom learning. *Computers & Education, 58*(1), 365-374.

Table 1
Sample Characteristics

Variable (value)	Mean (SD)		Frequency (%)	
Age	15.2	(0.6)		
Missing			1	(0.0)
Gender				
Male			1,586	(48.2)
Female			1,699	(51.7)
Missing			4	(0.1)
Ethnic origin				
Native			2,859	(87.2)
Other country			420	(12.8)
Missing			10	(0.3)
Mother educational level				
Low			748	(22.7)
Middle			1,548	(47.1)
High			854	(26.0)
Missing			139	(4.2)
Father educational level				
Low			1,000	(30.4)
Middle			1,345	(40.9)
High			769	(23.4)
Missing			175	(5.3)
Smartphone possession				
Yes			3,251	(98.8)
No			38	(1.2)
Missing			0	(0.0)
Italian language test score	210.1	(34.6)		
Missing			346	(10.5)
Math test score	212.5	(38.9)		
Missing			359	(10.9)
Random half-splitting procedure				
Subsample 1 (EFA)			1,644	(50.0)
Subsample 2 (CFA)			1,645	(50.0)

Note. $N = 3,289$, descriptive statistics of students' socio-demographic characteristics.

Table 2

Subsample 1: Item Descriptives and Exploratory Factor Loadings and Correlations (Model 3)

Items	M	SD	Skewness	Kurtosis	Factor1	Factor2	Factor3
					SPS	SAS	Residual covariance
SPS1	1.6	0.8	1.3	4.0	0.467	0.095	-0.110
SPS2	2.5	0.7	0.3	2.7	0.531	0.088	0.028
SPS3	2.6	0.8	0.1	2.4	0.630	-0.014	0.210
SPS4	1.9	1.1	1.0	2.5	0.569	0.023	-0.025
SPS5	2.7	1.1	-0.2	1.6	0.527	0.092	-0.023
SPS6	2.2	0.9	0.5	2.4	0.579	-0.069	0.057
SPS7	2.4	0.9	0.3	2.5	0.454	0.097	0.003
SAS1	2.0	1.3	1.1	3.5	-0.001	0.801	-0.065
SAS2	2.6	1.4	0.6	2.3	0.009	0.458	0.580
SAS3	2.9	1.5	0.3	2.2	-0.001	0.581	0.559
SAS4	2.8	1.4	0.3	2.2	0.140	0.617	-0.010
SAS5	2.7	1.4	0.5	2.4	0.054	0.557	0.212
SAS6	2.1	1.3	1.0	3.1	-0.089	0.536	0.095
SAS7	2.1	1.3	1.0	3.4	0.004	0.836	-0.098
SAS8	3.0	1.5	0.2	2.1	-0.041	0.526	0.368
SAS9	3.0	1.5	0.3	2.1	0.103	0.716	0.024
SAS10	2.7	1.5	0.5	2.4	0.089	0.636	-0.014
SPS-A					1.000		
SAS-SV					0.376	1.000	
Residual covariance					-0.003	0.096	1.000

Table 3

Summary of fit indices testing cross-group measurement invariance for Parents level of education, students' gender and ethnic origins

Group/Model	Fit Indices (MLMV estimation method)									
	χ^2	df	χ^2 p-val	RMSEA	CFI	TLI	McNCI	Δ CFI	Δ McNCI	$\Delta\chi^2$ p-val
Parents level of education										
Baseline Models										
Low (L)	218.2	112	0.000	0.046 [0.037 – 0.055]	0.921	0.904	-	-	-	-
Middle (M)	456.2	114	0.000	0.045 [0.040 – 0.048]	0.940	0.929	-	-	-	-
High (H)	426.3	114	0.000	0.049 [0.044 – 0.054]	0.923	0.908	-	-	-	-
Measurement Invariance										
M1: Configural	1089.5	340	0.000	0.046 [0.043 – 0.049]	0.931	0.917	0.887	-	-	-
M2: Metric	1118.6	370	0.000	0.044 [0.041 – 0.047]	0.931	0.924	0.887	0.000	0.000	0.658
M3: Residual Covariances	1121.0	378	0.000	0.043 [0.040 – 0.046]	0.932	0.926	0.888	0.001	0.001	0.550
M4: Scalar	1188.4	412	0.000	0.042 [0.040 – 0.045]	0.929	0.929	0.883	-0.003	-0.005	0.000
Students' gender										
Baseline Models										
Female	609.7	114	0.000	0.051 [0.047 – 0.055]	0.924	0.910	-	-	-	-
Male	502.9	114	0.000	0.047 [0.043 – 0.051]	0.929	0.915	-	-	-	-
Measurement Invariance										
M1: Configural	1108.3	228	0.000	0.049 [0.046 – 0.052]	0.926	0.912	0.872	-	-	-
M2: Metric	1184.4	243	0.000	0.049 [0.046 – 0.052]	0.921	0.912	0.864	-0.005	-0.008	0.000
M3: Residual Covariances	1192.1	247	0.000	0.049 [0.046 – 0.052]	0.921	0.913	0.863	0.000	-0.001	0.028
M4: Scalar	1397.2	264	0.000	0.052 [0.049 – 0.054]	0.905	0.902	0.838	-0.016	-0.025	0.000
M4.1: Scalar Partial SAS8-9	1312.1	262	0.000	0.050 [0.047 – 0.053]	0.912	0.909	0.850	-0.009	-0.013	0.000
Students' ethnic origins										
Baseline Models										
Natives	841.8	114	0.000	0.048 [0.045 – 0.051]	0.931	0.918	-	-	-	-
Migrants	226.4	114	0.000	0.050 [0.040 – 0.059]	0.920	0.904	-	-	-	-
Measurement Invariance										
M1: Configural	1030.9	228	0.000	0.047 [0.044 – 0.050]	0.929	0.915	0.883	-	-	-
M2: Metric	1047.8	243	0.000	0.045 [0.043 – 0.048]	0.929	0.920	0.882	0.000	-0.001	0.278
M3: Residual Covariances	1043.7	247	0.000	0.045 [0.042 – 0.048]	0.930	0.922	0.884	0.001	0.002	0.832
M4: Scalar	1108.9	264	0.000	0.045 [0.042 – 0.047]	0.925	0.923	0.879	-0.005	-0.005	0.000

Cut off values for fit indices. Close fit: $RMSEA \leq 0.05$; $CFI \geq 0.95$; $TLI \geq 0.95$. Acceptable fit: $RMSEA \leq 0.06$; $CFI \geq 0.90$; $TLI \geq 0.90$. Invariance test: $\Delta\chi^2 p\text{-value} \geq 0.050$; $\Delta CFI \leq 0.010$; $\Delta NCI \leq 0.020$.

Table 4
Cross-Group Latent Factor Means Comparison.

	SPS-A	SAS-SV
Parent educational level (Ref. Low)		
Middle	-0.231 (0.067)**	-0.083 (0.055)
High	-0.292 (0.067)**	-0.096 (0.059)
Ethnic origin (Ref. Native)		
Other	0.230 (0.064)**	0.105 (0.054)
Sex (Ref. Female)		
Male	-0.054 (0.040)	-0.207 (0.040)**
Sex (Ref. Female) partial		
Male	-0.054 (0.040)	-0.151 (0.041)**

Note. *p*-value: ** ≤ 0.01; * ≤ 0.05. Standardized estimates and standard errors within brackets.

Table 5
Results of the latent regression of Academic performance on problematic smartphone use

	TSI	TSM
SPS-A	-0.170 (0.030)**	-0.221 (0.029)**
SAS-SV	-0.031 (0.026)	-0.047 (0.026)
Parent educational level (Ref. Low)		
Middle	0.080 (0.025)**	0.087 (0.023)**
High	0.190 (0.025)**	-0.224 (0.025)**
Ethnic origin (Ref. Native)		
Other	-0.137 (0.018)**	-0.091 (0.017)**
Sex (Ref. Female)		
Male	-0.081 (0.018)**	0.132 (0.017)**

Note. *p*-value: ** ≤ 0.01; * ≤ 0.05. Abbreviations: smartphone pervasiveness scale (SPS-A); smartphone addiction scale (SAS-SV); INVALSI Italian language test score (TSI); INVALSI math test score (TSM).

Figure 1
Analytical Strategy

