Inman, Richard A.
Moreira, Paulo
Cunha, Diana
Castro, Jorge

Assessing the dimensionality of student school engagement survey : support for a multidimensional bifactor model

http://hdl.handle.net/11067/6356
https://doi.org/10.1016/j.psicoe.2020.03.001

Metadata

Issue Date 2020

Abstract The Student School Engagement Survey (SSES) is used to evaluate student engagement interventions run by the National Center for Student Engagement in the U.S. It was designed to measure the behavioral, emotional, and cognitive components of engagement, but its factorial structure has not been validated. To address this limitation, we tested the factorial structure of the Portuguese version of the SSES using a representative sample of 4,866 adolescents. An exploratory factor analysis revealed fiv...

Publisher Elsevier. Science Direct

Keywords Student engagement, Student school engagement survey, Bifactor model, Construct validity

Type article

Peer Reviewed yes

Assessing the Dimensionality of Student School Engagement Survey: Support for a multidimensional bifactor model

Richard A. Inman[2]

Paulo Moreira[1, 2]

Diana Cunha[2]

Jorge Castro[2]

[1]Instituto de Psicologia e de Ciências da Educação [*Institute of Psychology and Education*], Universidade Lusíada-Norte (Porto), Portugal
[2]Centro de Investigação em Psicologia para o Desenvolvimento (CIPD) [*The Psychology for Positive Development Research Center*], Portugal

Note. This document is the authors' version of the final accepted manuscript, published online in 27.4.2020 by *Revista de Psicodidáctica (English ed.)*.
doi: 10.1016/j.psicoe.2020.03.001

https://www.sciencedirect.com/science/article/pii/S253038052030006X

Corresponding author: Paulo Moreira, Instituto de Psicologia e de Ciências da Educação, Universidade Lusíada do Porto; Observatório da Melhoria e Eficácia da Escola; Centro de Investigação em Psicologia para o Desenvolvimento; Portugal. E-mail: paulomoreira@por.ulusiada.pt

Funding: Support for this research comes from national funds from the Fundação para a Ciência e Tecnologia I.P. (FCT) [*Portuguese Foundation for Science and Technology*], under the Projects PTDC/CPE-CED/122257/2010; PTDC/MHC-CED/2224/2014; and CIPD-BI-UID/PSI/04375/2016.

Abstract

The Student School Engagement Survey (SSES) is used to evaluate student engagement interventions run by the National Center for Student Engagement in the U.S. It was designed to measure the behavioral, emotional, and cognitive components of engagement, but its factorial structure has not been validated. To address this limitation, we tested the factorial structure of the Portuguese version of the SSES using a representative sample of 4,866 adolescents. An exploratory factor analysis revealed five theoretically meaningful factors describing subtypes of emotional and behavioral engagement, and teacher support for learning. A confirmatory factor analysis supported modelling a shortened version of the SSES with a bifactor model. Bifactor indices indicated total SSES scores are interpretable as a measure of a single student engagement construct. Finally, as evidence of concurrent validity, the scale had a strong positive correlation with an established measure of student engagement. The proposed version of the SSES is a psychometrically adequate measure of student engagement, although cannot be said to measure cognitive engagement.

Keywords: student engagement, Student School Engagement Survey, bifactor model, construct validity

Assessing the Dimensionality of the Student School Engagement Survey (Portuguese version): Support for a multidimensional bifactor model

A prevalent conceptualization in the research literature posits that student engagement has three distinct and dynamically interacting dimensions; behavioral, cognitive, and emotional engagement (Fredricks, Blumenfeld, & Paris, 2004; Jimerson, Campos, & Greif, 2003). A characteristic of these dimensions, as made evident by the continued debate over their content (Reschly & Christenson, 2012), is that they are broad constructs in their own rights. For example, behavioral engagement encompasses student conduct, participation in school activities, and tangible actions demonstrating willingness to overcome challenging material (Nguyen, Cannata, & Miller, 2016). Cognitive engagement is conceptualized as investment in learning (Yazzie-Mintz, 2007), perceptions and beliefs (Jimerson et al., 2003), and self-regulated learning (Appleton, Christenson, Kim, & Reschly, 2006). Finally, emotional engagement incorporates affective reactions to school (Connell & Wellborn, 1991), sense of belonging (Appleton et al., 2006), and the states relevant to student involvement such as interest (Skinner, Furrer, Marchand, & Kindermann, 2008).

An emerging theoretical framework suggests that the student engagement construct should also include students' perceptions of, and sense of relatedness with, relevant others as a way to recognize the goodness-of-fit between the student and their learning environment (Appleton et al., 2006; Christenson & Anderson, 2002; Reschly & Christenson, 2006; Sinclair, Christenson, Lehr, & Reschly, 2003). Indeed, Reschly and Christenson (2012) have argued that students' perceptions are the most accurate sources of information about objective reality. In line with this framework, a growing number of student engagement instruments now include items that capture students' perceptions of

support and relatedness with relevant others, including teachers, peers, and family members (e.g. Appleton et al., 2006).

The degree to which students are engaged with school has been shown to be an important predictor of academic achievement (Lee, 2014; Wang & Holcombe, 2010), rates of school dropout (Fall & Roberts, 2012; Wang & Fredricks, 2014), wellbeing (Wang, Chow, Hofkens, & Salmela-Aro, 2015), and prevalence of delinquency and mental illness (Li & Lerner, 2011). In turn, student engagement has been linked to background factors including socioeconomic status, gender, and race/ethnicity (Yazzie-Mintz, 2007), as well as educational factors such as prior academic performance (Chase, Hilliard, Geldhof, Warren, & Lerner, 2014; Moreira et al., 2018). Such findings imply that certain groups of students (e.g. those from lower socioeconomic families) are at risk of disengagement and its associated negative outcomes. Fortunately, a large body of research indicates that student engagement is responsive to teacher and school practices (Finn & Zimmer, 2012) meaning the prognoses of at-risk students can be improved by targeted interventions (Appleton, Christenson, & Furlong, 2008).

The Student School Engagement Survey

One organization involved in implementing student engagement interventions in the U.S. is the National Center for Student Engagement (NCSE); itself founded by the Colorado Foundation for Families and Children (CFFC). As a means to evaluate the effectiveness of its interventions (i.e. as an outcome measure of student engagement) the NCSE developed the Student School Engagement Survey (SSES; Finlay, 2006). According to Finlay (2006), the SSES was created by collaboratively selecting engagement items from multiple data sources, national surveys (National Longitudinal Study of Adolescent to Adult Health; ADD Health), journal articles (Fredricks et al., 2004), and the Core Measures (Center for Substance Abuse Prevention, 2003). The

chosen items were then grouped into three subscales representing the three major dimensions of student engagement as defined by Fredricks et al. (2004): cognitive, emotional, and behavioral.

A small amount of evidence from three pilot studies, presented in the initial NCSE research report (Finlay, 2006), suggests that the SSES may be a reliable measure. For example, across three independent samples Cronbach's alpha values for the SSES subscales ranged between $\alpha = .88$ to $\alpha = .90$ for emotional engagement, $\alpha = .87$ to $\alpha = .92$ for cognitive engagement, and $\alpha = .49$ to $\alpha = .80$ for behavioral engagement. Similar indications of scale reliability have been shown in published studies using student samples from Israel (Shoshani & Slone, 2013; $\alpha = .80$ to .91) and Mexico (Rodriguez & Boutakidis, 2013; $\alpha = .83$ and .89 for cognitive and emotional engagement scales), thus implying that this scale may have some cross-cultural validity. The NCSE pilot studies (Finlay, 2006) also provided evidence of convergent validity. Scores for the cognitive and behavioral scales had meaningful positive associations with students' grade point averages ($r = .37$ and $r = .35$ respectively), and in a sample of elementary school students, the emotional and cognitive scales were significantly correlated with Mathematics grades ($r = .48$ and $r = .40$), English grades ($r = .43$ and $r = .37$), and unexcused absences ($r = .61$ and $r = .46$).

Despite these initial findings, evidence supporting the psychometric adequacy of the SSES is limited. Remarkably, there has yet to be an empirical test of its construct validity using either exploratory factor analysis (EFA) or confirmatory factor analysis (CFA). Without an assessment of the factorial structure of the SSES, any conclusions made about the outcome of an intervention may be conceptually flawed. EFA in particular is required given the continued debate over the number and nature of engagement dimensions (Reschly & Christenson, 2012). Indeed, we argue that there are

several examples of items that do not match, in a theoretical sense, with their assigned dimensions. For example, the item “I study at home even when I don’t have a test” was included as an indicator of cognitive engagement. This would be more appropriately considered an aspect of behavioral engagement because it describes a tangible action aimed at learning academic content. A similar argument can be made for the item “I talk with people outside of school about what I am learning in class”, which was also classed under cognitive engagement. Furthermore, several items describing students’ perceptions of support and relatedness with teachers (e.g. “Most of my teachers care about how I’m doing”) were included as indicators of emotional engagement. Past research has strongly supported incorporating such perceptions as a distinct yet interrelated dimension of student engagement (Appleton et al., 2006; Moreira, Cunha, & Inman, 2019). Given these issues, there is an urgent need to identify the factorial structure of the SSES using EFA, and then to confirm this structure in an independent sample using CFA. Recent studies testing the dimensionality of other student engagement measures (e.g. Moreira et al., 2019) suggest that a bifactor model may be plausible for the SSES.

Assessing the Dimensionality of the SSES: The Bifactor Model

A recent body of work in the student engagement research literature has used the bifactor model approach to test the dimensionality of different student engagement instruments (Moreira et al., 2019; Wang, Fredricks, Ye, Hofkens, & Linn, 2016, 2017). Bifactor models are applicable when testing constructs that comprise multiple distinct-yet-related dimensions, and when there is a research interest in these specific dimensions in addition to the global construct (Chen, West, & Sousa, 2006). Bifactor models of student engagement are similar to second-order models in that they include a general student engagement factor. However, in the bifactor model, this factor accounts

for relationships between items rather than relationships between first-order factors (e.g. latent variables representing cognitive, emotional, and behavioral engagement). In addition, bifactor models also include specific factors (not mathematically or conceptually equal to first-order factors) that account for unique variance among groups of items beyond the general factor (Chen et al., 2006).

Recent psychometric studies of student engagement measures have utilized bifactor models to help disentangle the unique contributions of the multiple dimensions of engagement from the contribution of the global construct (Wang et al., 2016, 2017). In other words, bifactor models have been used to determine whether scale items are unidimensional and, therefore, whether they capture a theoretically unidimensional construct. A recent study by Moreira et al. (2019) supported modelling the Multifactorial Measure of Student Engagement as a bifactor model. Because this instrument included items that captured students' perceptions of support from teachers, peers, and family in addition to cognitive, emotional, and behavioral engagement, the findings implied that students' perceptions of support belong inside a global student engagement construct.

The Present Study

The SSES is a measure of student engagement largely used to evaluate the outcome of student engagement interventions in the U.S. However, there has yet to be an empirical test of this measure's construct validity using either EFA or CFA. This type of assessment is necessary because several of the SSES items do not align theoretically with their assigned dimensions. In the case of the present study, EFA was also necessary because we were using a Portuguese translation of the SSES in cultural setting distinct that for which the original English-language version was designed (i.e. Portugal). Hence, the primary objective of the study was to evaluate the dimensionality

of the Portuguese SSES using EFA, and then to confirm this structure in an independent sample via CFA. Because recent studies support modelling student engagement instruments using the bifactor model approach (Moreira et al., 2019), we aimed to test how well a bifactor model would represent the factor structure of the Portuguese SSES.

Method

Participants

This psychometric study uses participants from the first phase of a six-year longitudinal study into school effects on student engagement (for more details, see Moreira et al., 2018). In total, we used data from 4866 students attending 101 schools in Portugal. These students (54.7% female; 44.8% male) were enrolled in the seventh (n = 2247) or 10th grade (n = 2577) during this phase of data collection. Students in the seventh grade had a mean age of 12.5 years (SD = .75). Students in the 10th grade had a mean age of 15.6 years (SD = .90).

Measures

NCSE Student School Engagement Survey. Participants completed a version of the SSES that we had translated into European Portuguese. The authors of the study were granted permission to translate, adapt, and test the properties of this instrument by the Director of the NCSE. This instrument comprises 42 student engagement items in three subsections. Items 1 to 3 ("How important do you think…") are scored from 1 (*very important*) to 5 (*not at all important*). Items 4 to 28 ("How much do you agree with each of the following statements?") are scored from 1 (*strongly agree*) to 4 (*strongly disagree*). Items 29 to 42 ("How often are the following statements true for you?") are scored from 1 (*always*) to 4 (*never/almost never*).

We translated these items using the proposals of Mallinckrodt and Wang (2004). Firstly, a team of researchers who are fluent in both English and Portuguese prepared a

Portuguese translation of the original English version of the instrument. This first draft was then back-translated into English by a second team of researchers who are also fluent in both languages. This second team had no prior experience with the SSES. Experts in school engagement then determined the equivalence of these translated items and their reflection of either emotional, cognitive, or behavioral engagement. Finally, we asked a committee of peer consultants who were native Portuguese speakers and members of the target research population to examine the adapted scale using a "think aloud" procedure.

Student Engagement Instrument (SEI). Students also completed a second measure of student engagement: the brief 15-item Portuguese version of the SEI (Moreira & Dias, 2018). This instrument has two *cognitive engagement* subscales (control and relevance of schoolwork, and future aspirations and goals), and three subscales capturing students' perceptions of support from teachers, family, and peers respectively (collectively referred to as *affective engagement*; Betts et al., 2010). Items are scored from 1 (*totally disagree*) to 4 (*totally agree*). For the purpose of the present study, we calculated the mean of all 15 items to serve as a composite student engagement score. The psychometric properties of various versions of the SEI have been shown to be adequate in a large body of research (Appleton et al., 2006; Betts et al., 2010; Lovelace, Reschly, Appleton, & Lutz, 2014; Virtanen, Kiuru, Lerkkanen, Poikkeus, & Kuorelahti, 2016) including in Portuguese samples (Moreira & Dias, 2018; Moreira, Vaz, Dias, & Petracchi, 2009; Virtanen et al., 2018). In the present study, the reliability of the composite student engagement score was good ($\omega = .84$).

Procedures

Prior to collecting data, we obtained ethical approval from the ethics committee of Universidade de Lusíada, Portugal. Data collection happened at the start of the

academic year starting in 2013 (September – December 2013). In each participating school, a member of staff acted as a liaison between the school and the research team. This school representative planned the internal procedures for data collection. Questionnaires were administered to classes of students, gathered in a single room, under the supervision of the school representative. Note that as part of the broader longitudinal study students also completed several other measures, specifically several measuring subjective wellbeing.

Statistical Analysis

Analyses were conducted using R (R Core Team, 2019). The amount of missing data per item was small across all measures (< 1%). For the SSES, 5% of respondents had at least one missing data point, with 78% of these having missing data for just one or two items. For the SEI, 11% of respondents had at least one missing data point, with 89% of these having missing data for just one or two items.

The sample was divided randomly into two subsamples. The first sample (Sample A) was used for exploratory factor analysis (EFA; n = 2381). The second sample (Sample B) was used for confirmatory factor analysis (CFA; n = 2485). These subsamples did not differ in terms of mean age, $t(4854.5) = 0.06$, $p = .949$, or gender composition, $\chi^2(1) = 0.02$, $p = .890$. The full sample was used to test scale validity.

EFA. To determine the number of factors to extract we used parallel analysis (PA; Horn, 1965). A maximum likelihood analysis was then applied to test the factor solution proposed by PA. We used a direct oblimin rotation method because the dimensions of engagement were expected to be correlated. To optimize the Portuguese SSES, we chose to adopt a rule-of-thumb threshold of removing items with a factor loading < .40. Based on the recommendations of Raubenheimer (2004), we excluded

factors with fewer than two indicators with factor loadings ≥ .40. Missing values were imputed using the median.

CFA. Having identified an optimized factorial structure for the Portuguese SSES using EFA, we next used CFA to confirm this structure in an independent sample. Based on past research (Moreira et al., 2019), we chose to test a bifactor model. Because intraclass correlation coefficients (ICC) for the SSES items were low, indicating an average of 3.5% of variance in scores was explained by clustering at the school level, the CFA was conducted using the total covariance matrix. Because item scores were ordinal, we used a robust diagonally weighted least squares method (Li, 2016). Our judgment of model fit was guided by several indicators and heuristics for good fit: comparative fit index (CFI) ≥ .95 (Hu & Bentler, 1999), root mean square error approximation (RMSEA) ≤ .05 (Browne & Cudeck, 1992), and standardized root mean square residual (SRMR) < .05 (Hu & Bentler, 1999). Missing values were handled in this analysis using pairwise deletion.

Bifactor indices. First, we evaluated the reliability of the general factor by calculating omega (ω). Coefficient ω estimates the proportion of variance in the SSES total score attributed to all sources of variance. The same logic can be applied to calculate an index of reliability for each subscale: omegaS (ωs).

Next, we assessed the extent to which SSES total scores can be interpreted as a measure of a single construct despite multidimensionality in the scores. This was done by calculating omega hierarchical (ω_H), which represents the proportion of systematic variance in total scores accounted for by the general factor (Reise, Moore, & Haviland, 2010; Zinbarg, Yovel, Revelle, & McDonald, 2006). Values of ω_H larger than .75 indicate a total scale score can be interpreted as a measure of a single construct (Reise, Scheines, Widaman, & Haviland, 2013). The logic of ω_H can be applied to calculated

omega hierarchical for each subscale (ω_{HS}). These values represent the proportion of reliable variance of each subscale score after accounting for variability due to the general factor (Reise, Bonifay, & Haviland, 2013).

Next, we assessed the degree to which the multidimensional data are unidimensional. This was achieved by calculating the Explained Common Variance (ECV) index in conjunction with the Percentage of Uncontaminated Correlations (PUC) (Rodriguez, Reise, & Haviland, 2016). As a rule, bigger ECV values mean greater confidence in applying a unidimensional measurement model. Values of ECV larger than .70 suggest factor loadings from a unidimensional model are a good approximation of the factor loadings on a general factor from a bifactor model; i.e. an indication of less bias (Rodriguez et al., 2016). PUC, and indicator of model structure, moderates the association between ECV and model bias; when there are more uncontaminated correlations (when PUC is larger) relative bias can be low, even with smaller ECVs.

Finally, as a measure of construct replicability, we calculated the H index (Hancock & Mueller, 2001). Values of H > .80 imply a well-define latent construct that is likely to be stable across studies (Rodriguez et al., 2016).

Validity. A measure can be considered valid if it measures what it purports to measure (Borsboom, Mellenbergh, & Van Heerden, 2004). We tested the concurrent validity of the Portuguese SSES by assessing the association between the SSES and another validated measure of the student engagement construct: the SEI (Appleton et al., 2006). The SSES and SEI both measure a multidimensional student engagement construct with some shared dimensions (i.e. both measure students' perceptions of their relationships with teachers). Structural equation modelling (SEM) was used to estimate the relationship between the SSES bifactor structure and a composite indicator of student engagement from the SEI. Guided by the work of Yost and Finney (2018), SEI

student engagement was modeled as a latent factor with a single composite indicator. The unstandardized error variance of the composite indicator was calculated using the following equation: $(1-r_{xx}) \times var(x)$. r_{xx} corresponds to Cronbach's alpha for the composite score, and $var(x)$ is the variance. The latent factor representing the external variable was allowed to correlate with the general and specific factors. This approach provides an understanding of the relationship between external variables and the general factor isolated from the effects of specific factors.

Results

EFA

The PA of all 42 items supported retaining eleven factors (Figure 1). The standardized component loadings for the eleven-factor solution are presented in Table 1. The first theoretically meaningful factor had five items. The content of these items captured students' perceptions of teacher support for learning (example item: "The teachers at my school treat students fairly"). The second meaningful factor also had five items. These items captured student conduct (example item: "I complete my work on time"). The third meaningful factor had three items that captured study behaviors (example item: "I check my schoolwork for mistakes"). The fourth meaningful factor had three items that captured a sense of belonging and connectedness to school (example item: "I feel like I belong in my school"). The fifth meaningful factor had three items that captured' affective reactions to schoolwork (example item: "I feel excited by the work in school").

FIGURE 1 ABOUT HERE

Four factors were immediately excluded for having fewer than three items. Two further factors were also excluded. The first of these factors had items that also captured students' conduct. We chose to exclude this factor because it shared a very similar, and

thus redundant, item to the larger initial student conduct factor ("I follow rules in school" vs. "I follow the rules at school") and because it had fewer items. The items of the second factor had almost identical wordings ("When I first walked into my school I thought it was…"). We therefore chose to exclude this factor because it did not capture a theoretically meaningful dimensions of engagement.

TABLE 1 ABOUT HERE

CFA

The above described EFA suggested variance in the Portuguese SSES could be explained by five major factors with a total of 19 items: *teacher support for learning*, *student conduct*, *study behaviors*, *sense of belonging*, and *affective reactions to school and schoolwork*. The purpose of the CFA was to confirm this structure an independent sample. Based on past research (Moreira et al., 2019), we chose to test a bifactor model with a *general student engagement factor* and five specific factors corresponding to the five factors revealed by EFA (see Figure 2). Fitting the bifactor model resulted in the following fit: CFI = .957, RMSEA = .073, and SRMR = .059. Factor loadings for this model are shown in Table 2.

FIGURE 2 ABOUT HERE

Bifactor Indices

Bifactor indices are also presented in Table 2. Values for coefficients ω and ω_S (> .79) showed the Portuguese SSES scale and its subscales had good reliability. We assessed whether the SSES total score can be interpreted as a measure of a single construct, despite multidimensionality, by calculating ω_H. A general cut-off point for ω_H is .75 (Reise, Bonifay, et al., 2013). For this model, ω_H was .83. For all specific factors, the values for ω_{HS} were lower than for ω_H, which indicated much of the variance for each was attributed to the general factor (Rodriguez et al., 2016).

Values for ECV and PUC were .57 and .83 respectively. The value for ECV indicated that 57% of the common variance in items was attributable to the *general student engagement factor*; in other words, student engagement was the dominant construct being measured by the Portuguese SSES. Nonetheless, because ECV was below the threshold of .70 suggested by Rodriguez et al. (2016), researchers should model this scale using a bifactor approach for SEM and IRT analyses. However, the high value for PUC suggested that the relative bias of a unidimensional model may still be small, despite the lower ECV.

The value of H for the general factor was .91. This was greater than the recommended threshold of .80 (Rodriguez et al., 2016), implying a well-defined latent variable.

TABLE 2 ABOUT HERE

Concurrent Validity

SEM was used to test the factor correlation between the *general student engagement factor* of the bifactor model and a composite student engagement score from the SEI (Table 3). Consistent with what would be expected from two instruments measuring the same construct, the analysis revealed a strong positive correlation (r = .74, $p < .001$). The factor correlations between the composite student engagement score from the SEI and the specific factors of the bifactor model were weak ($rs < .20$), but nonetheless statistically significant for *sense of belonging*, *affective reactions to school and schoolwork*, and *student conduct*. This finding implies that the observed correlation between the SSES and SEI would have been overestimated if we had used a unidimensional model of the SSES.

TABLE 3 ABOUT HERE

Discussion

The SSES is an assessment instrument originally designed for use by a national organization in the U.S. (the NCSE) to evaluate the effectiveness of student engagement interventions. To be used for this purpose, the SSES should have adequate psychometric properties; that is, it should measure student engagement in a reliable and valid way. Because evidence to support this is currently limited, the broad objective of the study was to test the psychometric properties of the SSES. More specifically, the factorial structure of this instrument has never been tested, which is important because there are multiple instances where items have a theoretical mismatch with their assigned subscales. To address this issue, the current paper was largely dedicated to testing the dimensionality of the Portuguese SSES using exploratory and confirmatory factor analysis.

The EFA indicated that 19 of the original 42 items could be grouped into five theoretically meaningful factors. Two of these factors corresponded to two of the different domains of emotional engagement highlighted by Fredricks, Blumenfeld, and Paris (2004): *affective reactions to school and schoolwork* and *sense of belonging*. Two further factors represented different domains of behavioral engagement: *student conduct* and *study behaviors*. These subdomains are consistent with theoretical distinctions made in research specific to behavioral engagement (Nguyen et al., 2016), and were consistent with factors identified in other student engagement instruments, including the MMSE. The final factor, *teacher support for learning,* included five items that captured students' thoughts about the quality of support received directly or indirectly from teachers (e.g. "The teachers at my school treat students fairly", "The discipline at my school is fair", "Most of my teachers care about how I'm doing", "I am getting a good education at my school"), or the quality of student-teacher relationships ("I like most of my teachers at school"). Such items had been classified by the NCSE as indicators of

cognitive engagement despite themselves having defined this aspect of engagement as "psychological investment in learning, a desire to go beyond the requirements and a preference for challenge" (Finlay, 2006, p.3). Instead, we propose that such items are consistent with students' perceptions of contextual influences (in this instance the influence from teachers) as captured by the affective engagement dimension of the SEI (Appleton et al., 2006) and the MMSE (Moreira et al., 2019).

A major finding of the study was that the shorted version of the SSES (with the 19 items retained after EFA) could be modelled as a general student engagement factor after partialling out shared common variance from five specific factors, i.e. via a bifactor model. High values for omega hierarchical and an ECV greater than .50 suggested that the majority of variance in SSES scores was explained by the general student engagement factor. These results therefore add to a growing body of research that demonstrates student engagement with school is a multidimensional construct, and should be modelled as a bifactor model for use with SEM or IRT, but that total scores across these dimensions are also interpretable as an indicator of a single higher-order construct (Moreira et al., 2019; Wang et al., 2017). Our study also builds on evidence that supports engagement frameworks that include students' subjective perceptions of support from relevant others belong within the global student engagement construct (Appleton et al., 2006; Moreira et al., 2019). In sum, having demonstrated that the internal structure of the SSES is consistent with the structure of the student engagement construct as proposed by recent frameworks and empirical studies (e.g. Moreira et al., 2019), the study indicates that the SSES has structural validity (Messick, 1995).

A further contribution of the current study was the demonstration that the shortened 19-item SSES had concurrent validity. A scale can be said to have validity when it measures what it purports to measure (Borsboom et al., 2004). This was tested

by assessing the factor correlations between a composite score from a benchmark measure of student engagement, in this case the well-validated SEI (Appleton et al., 2006; Betts et al., 2010; Moreira & Dias, 2018), and the general student engagement factor of the SSES bifactor model. As anticipated, the association between these two measures (r = .74) can be described as being moderate (Ferguson, 2009) to strong (Cohen, 1988). This finding thus indicates the SSES general student engagement factor shares a strong conceptual similarity with the student engagement construct assessed by the SEI. This finding is noteworthy because both instruments measure different dimensions of student engagement. For example, the SEI captures cognitive engagement (future aspirations and goals; perceptions of control and relevance of schoolwork) and students perceptions of support from peers and family, while the SSES does not. In turn, the SSES captures aspects of behavioral engagement (student conduct and study behaviors) and affective reactions to school, while the SEI does not. The strong correlation between these two measures therefore suggests that they capture a shared global student engagement construct, but that each covers some unique conceptual space.

Implications for Practice

It is important to have psychometrically sound measures of student engagement, themselves based on a theoretically accurate conceptualization of the student engagement construct, to inform interventions and address important educational issues such as disengagement and improving school attendance. Although the SSES is currently used by the NCSE to evaluate the effectiveness of student engagement interventions, the present study is the first to test this instrument's factorial structure via factor analysis. Because the results of the EFA revealed a different factorial structure to that proposed by the NCSE, the first implication of this study is that users of the SSES

should carefully consider how its items relate to theoretical dimensions. Fortunately, the present study provides an insight into the dimensions of student engagement captured by the SSES. The authors propose that a shortened 19-item version of the SSES may be a useful tool for measuring a global student engagement construct, despite its omission of several relevant dimensions (i.e. cognitive engagement and students' perceptions of support from peers and family).

Study Strengths and Limitations

A strength of this study was that it was conducted using a large sample of seventh and 10th graders (with an *n* close to 5000 individuals) from over 100 schools in Portugal. Moreover, this school sample captured the diverse nature of schools in Portugal, including private and public institutions; middle (seventh to ninth grades), secondary (10th to 12th grades), and mixed (seventh to 12th grades) schools; as well as schools from urban and rural communities. In short, the sample used in the present sample was considered to be representative of seventh and 10th graders in Portugal. In addition, although power calculations were not made, the sample size far exceeded all rules of thumb for EFA and CFA (Kyriazos, 2018), implying that the study had adequate statistical power. Nonetheless, some specific characteristics of the study sample suggest that future work is needed to determine if the SSES is valid in and across different groups of students. For example, studies should test whether the SSES has measurement equivalence across different age groups (including primary school students). Finally, it is important to recognize that the results were based on a Portuguese-language instrument applied in Portugal. Further research is required to determine whether the SSES has cross-cultural measurement equivalence.

References

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with

school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, *45*(5), 369–386. https://doi.org/10.1002/pits

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology*, *44*(5), 427–445. https://doi.org/10.1016/j.jsp.2006.04.002

Betts, J. E., Appleton, J. J., Reschly, A. L., Christenson, S. L., & Huebner, S. E. (2010). A study of the factorial invariance of the Student Engagement Instrument (SEI): Results from middle and high school students. *School Psychology Quarterly*, *25*(2), 84–93. https://doi.org/10.1037/a0020259

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*(4), 1061–1071. https://doi.org/10.1037/0033-295X.111.4.1061

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230–258. https://doi.org/10.1177/0049124192021002005

Center for Substance Abuse Prevention. (2003). *Core measures initiative phase I recommendations*. Center for Substance Abuse Prevention.

Chase, P. A., Hilliard, L. J., Geldhof, G. J., Warren, D. J. A., & Lerner, R. M. (2014). Academic achievement in the high school years: The changing role of school engagement. *Journal of Youth and Adolescence*, *43*(6), 884–896. https://doi.org/10.1007/s10964-013-0085-4

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and secondary-order models of quality of life. *Multivariate Behavioral Research*, *41*(2), 189–225. https://doi.org/10.1207/s15327906mbr4102

Christenson, S. L., & Anderson, A. R. (2002). Commentary: The centrality of the learning context for students' academic enabler skills. *Journal of School Psychology, 31*(3), 2293–2778.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Eds.), *Minnesota symposium on child psychology* (pp. 43–77). Hillsdale, NJ: Erlbaum.

Fall, A. M., & Roberts, G. (2012). High school dropouts: Interactions between social context, self-perceptions, school engagement, and student dropout. *Journal of Adolescence*, *35*(4), 787–798. https://doi.org/10.1016/j.adolescence.2011.11.004

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, *40*(5), 532–538. https://doi.org/10.1037/a0015808

Finlay, K. (2006). *Quantifying school engagement: Research report*. Retrieved from http://schoolengagement.org/school-engagement-services/ncse-publications/reports/

Finn, J. D., & Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter? In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 97–131). New York, NY: Springer.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent

variable systems. In R. Cudecl, S. du Toit, & D. Sörbom (Eds.), *Structural Equation Modeling: Present und Future—A Festschrift in Honor of Karl Joreskog* (pp. 195–216). Lincolnwood, IL: Scientific Software International.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrikaa, 30*(2), 179–185. https://doi.org/10.1007/BF02289447

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

Jimerson, S. R., Campos, E., & Greif, J. L. (2003). Toward an understanding of definitions and measures of school engagement and related terms. *The California School Psychologist*, *8*(1), 7–27. https://doi.org/10.1007/BF03340896

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology, 9*, 2207–2230. https://doi.org/10.4236/psych.2018.98126

Lee, J.-S. (2014). The relationship between student engagement and academic performance: Is it a myth or reality? *Journal of Educational Research, 107*(3), 177–185. https://doi.org/10.1080/00220671.2013.807491

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Li, Y., & Lerner, R. M. (2011). Trajectories of school engagement during adolescence: Implications for grades, depression, delinquency, and substance use. *Developmental Psychology*, *47*(1), 233–247. https://doi.org/10.1037/a0021307

Lovelace, M. D., Reschly, A. L., Appleton, J. J., & Lutz, M. E. (2014). Concurrent and predictive validity of the student engagement instrument. *Journal of*

Psychoeducational Assessment, *32*(6), 509–520. https://doi.org/10.1177/0734282914527548

Mallinckrodt, B., & Wang, C. D. (2004). Quantitative methods for verifying semantic equivalence of translated research instruments: A Chinese version of the Experiences in Close Relationships Scale. *Journal of Counseling Psychology*, *51*, 368–379. https://doi.org/10.1037/0022-0167.51.3.368

Messick, S. (1995). Standards of Validity and the Validity of Standards in Performance Asessment. *Educational Measurement: Issues and Practice*, *14*(4), 5–8. https://doi.org/10.1111/j.1745-3992.1995.tb00881.x

Moreira, P., Cunha, D., & Inman, R. A. (2019). An integration of multiple student engagement dimensions into a single measure and validity-based studies. *Journal of Psychoeducational Assessment*, Advance online publication. https://doi.org/10.1177/0734282919870973

Moreira, P., Dias, A., Matias, C., Castro, J., Gaspar, T., & Oliveira, J. (2018). School effects on students' engagement with school: Academic performance moderates the effect of school support for learning on students' engagement. *Learning and Individual Differences*, *67*, 67–77. https://doi.org/10.1016/j.lindif.2018.07.007

Moreira, P., & Dias, M. A. (2018). Tests of factorial structure and measurement invariance for the Student Engagement Instrument: Evidence from middle and high school students in Portugal. *International Journal of School & Educational Psychology*, *7*(3), 174–186. https://doi.org/10.1080/21683603.2017.1414004

Moreira, P., Vaz, F. M., Dias, P. C., & Petracchi, P. (2009). Psychometric properties of the portuguese version of the student engagement instrument. *Canadian Journal of School Psychology*, *24*(4), 303–317. https://doi.org/10.1177/0829573509346680

Nguyen, T. D., Cannata, M., & Miller, J. (2016). Understanding student behavioral

engagement: Importance of student interaction with peers and teachers. *Journal of Educational Research*, *111*(2), 163–174. https://doi.org/10.1080/00220671.2016.1220359

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Core Team.

Raubenheimer, J. (2004). An item selection procedure to maximise scale reliability and validity. *SA Journal of Industrial Psychology*, *30*(4), 59–64. https://doi.org/10.4102/sajip.v30i4.168

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, *95*(2), 129–140. https://doi.org/10.1080/00223891.2012.725437

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, *92*(6), 544–559. https://doi.org/10.1080/00223891.2010.496477

Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2013). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. *Educational and Psychological Measurement*, *73*(1), 5–26. https://doi.org/10.1177/0013164412449831

Reschly, A. L., & Christenson, S. L. (2006). Prediction of dropout among students with mild disabilities: A case for the inclusion of student engagement variables. *Remedial and Special Education*, *27*(5), 276–292. https://doi.org/10.1177/07419325060270050301

Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness:

Evolution and future directions of the engagement construct. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 3–19). https://doi.org/10.1007/978-1-4614-2018-7

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment, 98*(3), 223–237. https://doi.org/10.1080/00223891.2015.1089249

Rodriguez, J. L., & Boutakidis, I. P. (2013). The association between school engagement and achievement across three generations of Mexican American students. *Journal of the Association of Mexican American Educators, 7*(1), 5–12.

Shoshani, A., & Slone, M. (2013). Middle school transition from the strengths perspective: Young adolescents' character strengths, subjective well-being, and school adjustment. *Journal of Happiness Studies, 14*(4), 1163–1181. https://doi.org/10.1007/s10902-012-9374-y

Sinclair, M. F., Christenson, S. L., Lehr, C. A., & Reschly, A. L. (2003). Facilitating student engagement: Lessons learned from Check & Connect longitudinal studies. *The California School Psychologist, 8*, 29–41. https://doi.org/10.1007/BF03340894

Skinner, E. A., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology, 100*(4), 765–781. https://doi.org/10.1037/a0012840

Virtanen, T. E., Kiuru, N., Lerkkanen, M.-K., Poikkeus, A.-M., & Kuorelahti, M. (2016). Assessment of student engagement among junior high school students and associations with self-esteem, burnout, and academic achievement. *Journal for Educational Research Online, 8*(2). https://doi.org/10.1177/1741143214558576

Virtanen, T. E., Moreira, P., Ulvseth, H., Andersson, H., Tetler, S., & Kuorelahti, M. (2018). Analyzing measurement invariance of the Students' Engagement

Instrument brief version: The cases of Denmark, Finland, and Portugal. *Canadian Journal of School Psychology*, *33*(4), 297–313. https://doi.org/10.1177/0829573517699333

Wang, M.-T., Chow, A., Hofkens, T., & Salmela-Aro, K. (2015). The trajectories of student emotional engagement and school burnout with academic and psychological development: Findings from Finnish adolescents. *Learning and Instruction*, *36*, 57–65. https://doi.org/10.1016/j.learninstruc.2014.11.004

Wang, M.-T., & Fredricks, J. A. (2014). The reciprocal links between school engagement, youth problem behaviors, and school dropout during adolescence. *Child Development*, *85*(2), 722–737. https://doi.org/10.1111/cdev.12138

Wang, M.-T., Fredricks, J. A., Ye, F., Hofkens, T. L., & Linn, J. S. (2016). The Math and Science Engagement scales: Scale development, validation, and psychometric properties. *Learning and Instruction*, *43*, 16–26. https://doi.org/10.1016/j.learninstruc.2016.01.008

Wang, M.-T., Fredricks, J. A., Ye, F., Hofkens, T., & Linn, J. S. (2017). Conceptualization and assessment of adolescents' engagement and disengagement in school: A multidimensional school engagement scale. *European Journal of Psychological Assessment*, *Advance on*. https://doi.org/10.1027/1015-5759/a000431

Wang, M.-T., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal*, *47*(3), 633–662. https://doi.org/10.3102/0002831209361209

Yazzie-Mintz, E. (2007). *Voices of students on engagement: A report on the 2006 high school survey of student engagement*. Retrieved from https://pdfs.semanticscholar.org/504f/ce0c038d6cf1206d67566d54ebe1df8d0d2d.p

df?_ga=2.246060676.10773328.1579168923-334591954.1579168923

Yost, A. B., & Finney, S. J. (2018). Assessing the unidimensionality of trait reactance using a multifaceted model assessment approach. *Journal of Personality Assessment, 100*(2), 186–196. https://doi.org/10.1080/00223891.2017.1280044

Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for ωh. *Applied Psychological Measurement*, *30*(2), 121–144. https://doi.org/10.1177/0146621605278814

Table 1

Item factor loadings following an exploratory factor analysis (EFA) with oblimin rotation using Sample A (n = 2381).

Item	Factor											Item Text (English original)
	1	2	3	4	5	6	7	8	9	10	11	
13	1.00											Most of my classes are boring. (R)
11	1.00											I will fail no matter how hard I try (R)
7		.64										**The teachers at my school treat students fairly.**
12		.56										**The discipline at my school is fair.**
9		.55										**I like most of my teachers at school.**
10		.54										**I am getting a good education at my school.**
14		.42										**Most of my teachers care about how I'm doing.**
8												I feel safe in my school.
15												I learn a lot from my classes.
24			.66									**I treat my classmates with respect.**
26			.66									**I treat my teachers with respect.**
28			.49									**I follow rules in school.**
25			.46									**I complete my work on time.**
23			.41									**I come to class prepared**
27												I try my best on homework.
21				.88								When I first walked into my school I thought it was friendly.
20				.77								When I first walked into my school I thought it was good.
22				.52								When I first walked into my school I thought it was clean.
36					.68							**I check my schoolwork for mistakes.**
34					.63							**I study at home even when I don't have a test.**
35					.48							**I talk with people outside of school about what I am learning in class.**
37												If I don't understand what I read, I go back and read it over again.

5	.77					**I feel like I belong in my school.**
4	.57					**I feel close to people at my school.**
6	.54					**I am happy to be at my school.**
32		.67				**I am interested in the work I get to do in my classes.**
31		.63				**I feel excited by the work in school.**
41		.40				**I enjoy the work I do in class.**
33						My classroom is a fun place to be.
29			.56			I follow the rules at school.
30			.49			I get in trouble at school. (R)
17			.48			I respect most of my teachers.
18						School is a waste of my time. (R)
42				.63		I feel I can go to my teacher(s) with the things that I need to talk about.
16						There is an adult at school that I can talk to about my problems.
19						Most of my teachers understand me.
38						Most of my teachers praise me when I work hard.
2					.57	The things you are learning in school are going to be to you later in life?
1					.56	It is to get good grades?
3						It is to attend school every day?
40					.50	I get good grades in school.
39					.44	I try my best at school.

Notes. (R) = item was reverse coded. Items in bold are items tested in CFA. Only factor loadings > .40 reported.

Table 2.

Fully-standardized factor loadings and residual variances for the Student School Engagement Survey items based on the bifactor model, tested using Sample B (n = 2485).

Item	λ_{GEN}	$\lambda_{EMO.B}$	λ_{TSL}	$\lambda_{BEH.SC}$	$\lambda_{EMO.AR}$	$\lambda_{BEH.SB}$	Residual Variance
4	.445	.476					.575
5	.548	.812					.040
6	.622	.409					.446
7	.579		.441				.470
9	.595		.457				.437
10	.697		.368				.379
12	.548		.434				.511
14	.497		.403				.591
23	.677			.296			.454
24	.591			.586			.307
25	.594			.344			.529
26	.607			.677			.173
28	.580			.475			.438
31	.602				.502		.386
32	.622				.619		.230
41	.677				.348		.421
34	.468					.610	.409
35	.449					.445	.600
36	.565					.596	.326
ω	.947						
ω_S		.838	.844	.887	.848	.787	
ω_H	.825						
ω_{HS}		.440	.289	.336	.317	.436	
ECV	.569						
ECV_S		.543	.341	.398	.385	.556	
H	.910	.708	.521	.654	.523	.582	

Notes.
GEN = general student engagement factor; EMO.B = emotional engagement: belonging; TSL = teacher support for learning; BEH:SC = behavioral engagement: student conduct: EMO.AR = emotional engagement: affective reactions; BEH.SB = behavioral engagement: study behaviors;

ω = Omega; ECV = Explained Common Variance; H = Construct Replicability Index (Hancock & Mueller, 2001)

Table 3

Factor correlations between SSES factors from the bifactor model and composite student engagement score from the SEI. Analysis conducted with the full sample (n = 4866).

	Composite Student Engagement (SEI)	
	r	*p*
General factor		
Student Engagement (SSES)	.740	<.001
Specific factors		
Emotional engagement: belonging	.051	.061
Emotional engagement: affective reactions	.139	<.001
Behavioral engagement: student conduct	.139	.001
Behavioral engagement: study behaviors	.004	.914
Teacher support for learning	.038	.182

Note. SSES = Student School Engagement Survey; SEI = Student Engagement Instrument.

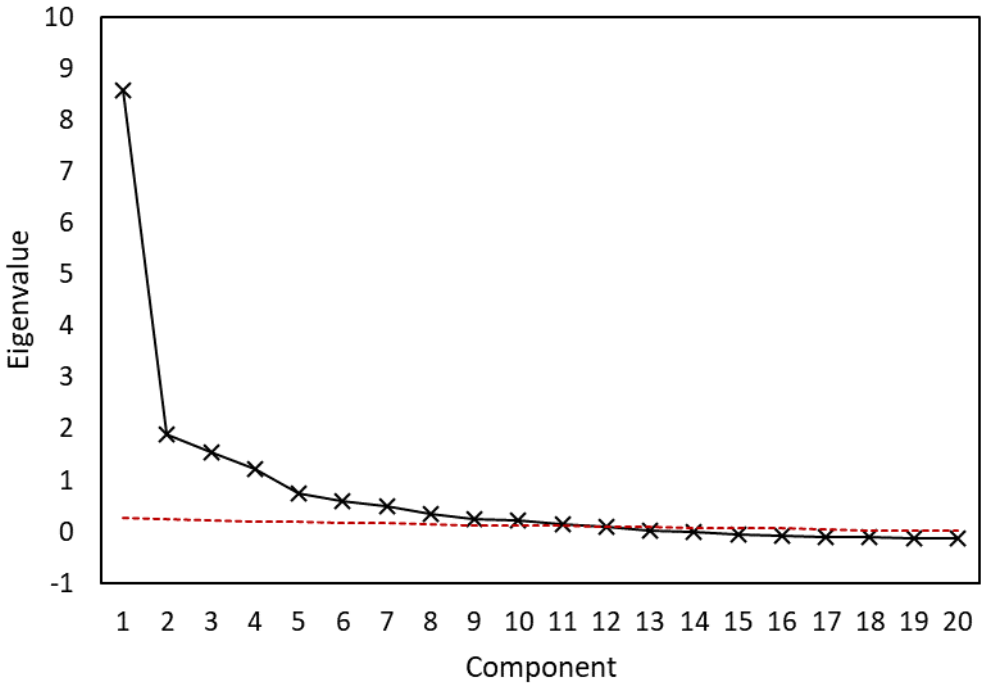

Figure 1. Scree plot of actual (solid line with markers) versus simulated eigenvalues (dotted line).

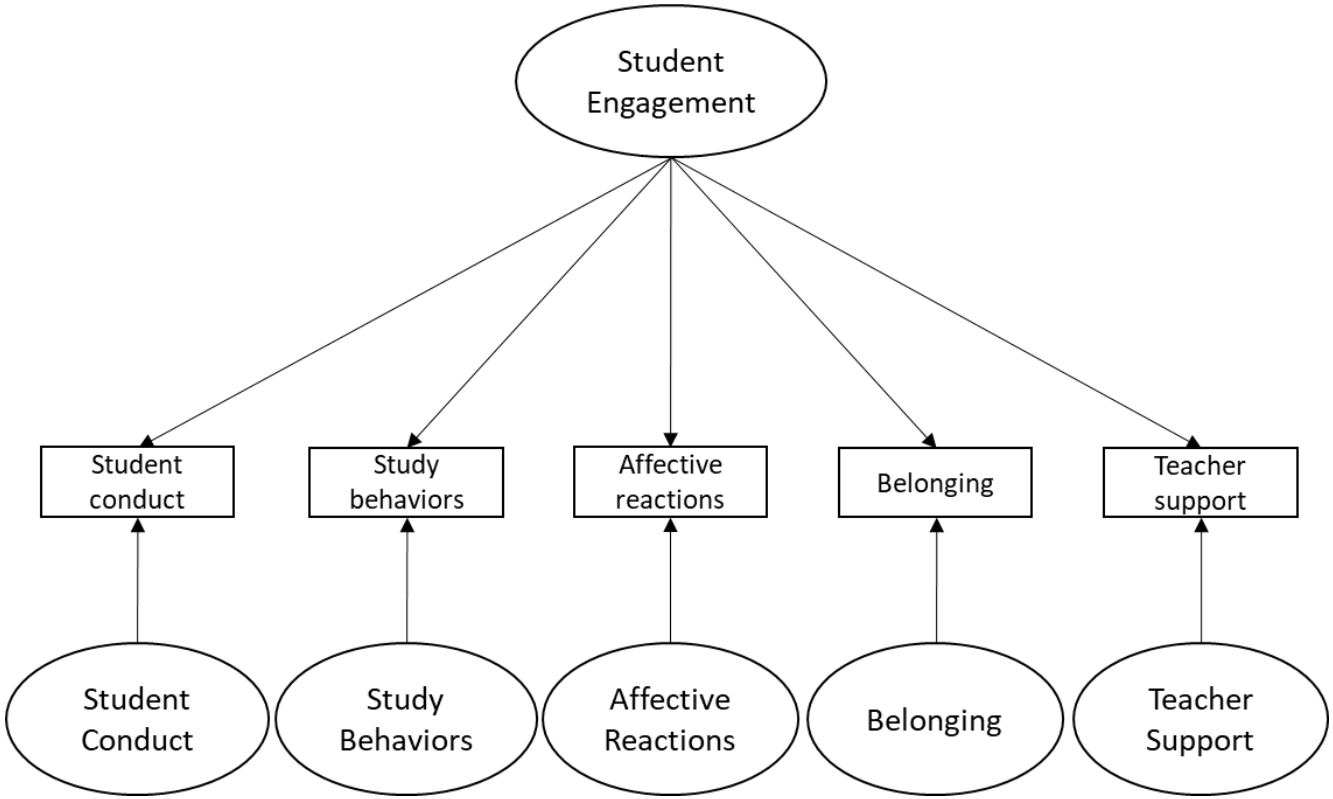

Figure 2. Bifactor model tested in CFA. Ellipses represent general and specific factors. Rectangles represent collections of items.