Open access • Journal Article • DOI:10.1007/S12144-018-0082-6

University Student Engagement Inventory (USEI): Psychometric properties

Jorge Sinval, Joana R. Casanova, João Maroco, Leandro S. Almeida

Institutions: International Sleep Products Association, University of Minho

Published on: 01 Apr 2021 - Current Psychology (Springer US)

Topics: Student engagement, Measurement invariance, Confirmatory factor analysis and Graduation

University Student Engagement Inventory (USEI): Psychometric properties

Jorge Sinval[1], Joana R. Casanova[2], João Marôco[1], & Leandro S. Almeida[2]

1 William James Center for Research, ISPA – Instituto Universitário

2 Research Centre in Education (CIEd), Institute of Education, University of Minho

Acknowledgments

Jorge Sinval received funding from the William James Center for Research, Portuguese Science Foundation (FCT UID/PSI/04810 /2013). Leandro S. Almeida and Joana R. Casanova received funding from CIEd – Research Centre on Education, projects UID/CED/1661/2013 and UID/CED/1661/2016, Institute of Education, University of Minho, through national funds of FCT/MCTES-PT. Joana R. Casanova received funding from the Portuguese Science Foundation (FCT) as a Doctoral Grant, under grant agreement number SFRH/BD/117902/2016.

Contact author

E-mail: jpmaroco@ispa.pt

The final publication is available at link.springer.com (http://dx.doi.org/10.1007/s12144-018-0082-6).

University Student Engagement Inventory (USEI): Psychometric properties

Abstract

Academic engagement describes students' investment in academic learning and achievement and is an important indicator of students' adjustment to university life, particularly in the first year. A tridimensional conceptualization of academic engagement has been accepted (behavioral, emotional and cognitive dimensions). This paper tests the dimensionality, internal consistency reliability and invariance of the University Student Engagement Inventory (USEI) taking into consideration both gender and the scientific area of graduation. A sample of 908 Portuguese first-year university students was considered. Good evidence of reliability has been obtained with ordinal alpha and omega values. Confirmatory factor analysis substantiates the theoretical dimensionality proposed (second-order latent factor), internal consistency reliability evidence indicates good values and the results suggest measurement invariance across gender and the area of graduation. The present study enhances the role of the USEI regarding the lack of consensus on the dimensionality and constructs delimitation of academic engagement.

Keywords: academic engagement, higher education, first-year students, assessment, measurement invariance

Introduction

Research on engagement emerged from professional and occupational contexts. In these contexts, engagement is defined as a positive psychological state that is characterized by vigor, dedication and absorption associated with work-related well-being (Bakker et al. 2008; Hirschi 2012; Schaufeli and Bakker 2010). In recent years engagement has also been studied in educational contexts, namely in higher education (Bresó et al. 2011; Christenson and Reschly 2010; Kuh 2009; Vasalampi et al. 2009). These studies are often present in international research concerning academic learning and achievement (Krause and Coates 2008; Schaufeli, Martinez, Marques Pinto, et al. 2002).

Students' academic engagement can be defined as the time, intention and energy students devote to educationally sound activities. Academic engagement is related to the policies and practices that institutions use to induce students to take part in those activities (Hodson and Thomas 2003; Kuh 2005; Wierstra et al. 2003). Research has established that engaged students invest more in their performance, participate more and tend to develop mechanisms to help them persist and self-regulate their learning and achievement (Klem and Connell 2004; National Research Council and Institute of Medicine 2004).

Academic engagement is associated with a positive way of experiencing academic activities and contexts, since it is related to positive academic and social outcomes (Klem and Connell 2004; Wonglorsaichon et al. 2014), to satisfaction and self-efficacy (Coetzee and Oosthuizen 2012), and to a reduction of achievement problems, burnout and dropout (Chapman et al. 2011; Christenson et al. 2012; Christenson and Reschly 2010; Eccles and Wang 2012; Elmore and Huebner 2010; Finn and Zimmer 2012; Fredricks et al. 2004, 2011; Gilardi and Guglielmetti 2011; Reschly and Christenson 2012a).

As engagement is a broad meta-construct it can be problematic because various definitions exist both within and across the different types of engagement (Fredricks et al. 2016). Two dominant conceptualizations of academic engagement have emerged in the literature (for a recent debate on academic engagement see Senior and Howard 2015). Schaufeli, Martinez, Marques-Pinto, Salanova and Bakker (2002) adapted the Utrecht Work Engagement Scale (UWES) from the business organizations' perspective to measure student engagement in university settings. The adapted scale, the UWES – Student version (UWES-S), uses the same three work engagement dimensions (vigor, absorption and dedication) adapted to the university context by rephrasing some of the original UWES items. The other predominant student academic engagement conceptualization by Fredricks et al. (2004) defines academic engagement as a multidimensional construct, integrating behavioral, emotional and cognitive dimensions, which is usually in line with the notion that the behavioral component corresponds to vigor, the emotional one to dedication and the cognitive one to absorption (Christensen 2017). However, criticisms have been raised regarding Salanova et al.'s (2002) and Fredricks et al.'s (2004) student academic engagement conceptualizations. The former was a simple adaptation of the workplace to the university context; the latter was derived mainly for high school students (Marôco et al. 2016).Theorizing academic engagement as a multidimensional construct allows for the better generalization and understanding of academic engagement as a combination of its several factors. Also, analysis of the engagement first-order factors (behavioral, emotional and cognitive) allows for pinpointing the different contribution to overall engagement and direct interventions.

Clarification is needed since some theoretical frameworks almost overlap with previous literature (Fredricks 2015). In the academic engagement literature, there is a need

for clear definitions with differentiation between the dimensions within the adopted framework (Fredricks et al. 2004). Raising the importance of having measures that take this into consideration without crossing the content of different dimensions of different factors increases the utility of analyzing the validity evidence of multidimensional psychometric instruments. Marôco, Marôco, Campos and Fredricks (2016) reviewed the main criticisms of both approaches and developed the University Student Engagement Inventory (USEI). This inventory includes the behavioral, cognitive and emotional dimensions of academic engagement, which is the definition and division of dimensions adopted by most research (Fredricks 2015). The behavioral dimension is related to behaviors such as attending classes, arriving on time, doing prescribed tasks/homework in scheduled time, participating in activities in and out of the classroom, and respecting the social and institutional rules. The cognitive dimension refers to all the students' thoughts, perceptions and strategies related with the acquisition of knowledge or development of competencies to academic activities, for example their study methods, learning approaches and academic self-regulation. The emotional dimension refers to positive and negative feelings and emotions related to the learning process, class activities, peers and teachers, for example a sense of belonging, enthusiasm, and motivation (Antúnez et al. 2017; Carter et al. 2012; Marôco et al. 2016; Sheppard 2011). Validity evidence based on response processes (i.e. face validity) of the behavioral, cognitive and emotional as dimensions of academic engagement was evaluated by a focus group of university students and psychologists in the original proposal of Mâroco et al. (2016). In this study, we focus on the validity evidence based on the USEI's internal structure.

Although there is a consensus about the relevance of this construct to the explanation of academic behavior and learning, there is not a precise delimitation of the construct and its

dimensionality (Christenson et al. 2012; Fredricks and McColskey 2012; Kahu 2013; Reschly and Christenson 2012; Wolf-Wendel et al. 2009). A debate is still ongoing concerning the definition and internal structure of the academic engagement construct. This conceptual haziness (Appleton et al. 2008) extends to the dimensionality of the construct's instruments: Some authors assume it to be a unidimensional primary factor or a second-order factor as it is a general motivational trait or state, while other authors defend its multidimensionality, but without consensus regarding the number of dimensions (Fredricks et al. 2004; Handelsman et al. 2005; Lin and Huang 2018; Reschly and Christenson 2012b).

In this paper, we focus on the USEI for the university context and evaluate one of the sources of evidence proposed in the *Standards for Educational and Psychological Testing* (American Educational Research Association et al. 2014) regarding the validity evidence based on the internal structure. Specifically, we aim to find good validity evidence regarding the dimensionality of the first-order three-factor model (H1) of a possible second-order latent factor model (H2), measurement invariance for gender (H3) and for the scientific area of college graduation (H4), and good evidence of reliability of the scores through internal consistency using several estimates (H5). This type of validity indicators intends to demonstrate the relevance of an instrument that simultaneously can be useful to investigation and practice. Namely, shown evidence of a meta-construct (academic engagement) which is useful for research, demonstrating the utility of its specific domains for interventions with specific students´ subgroups.

Method

Validity is a vital issue when it refers to the quality of psychometric scales, and it refers to the extent to which the evidence supports the interpretation of scale scores (Crutzen

and Peters 2017). Validity concerns the understanding of scale scores in a specific study; it isn't a characteristic of a scale in itself (American Educational Research Association et al. 2014). Consequently, evidence from other studies must be used to justify the choice of a specific scale, although in a strict sense it doesn't guarantee the same validity evidence in a new study (Crutzen and Peters 2017). Nevertheless, every study that uses psychometric scales must pay attention to the validity evidence brought by each scale in each study. Historically, different types of validity have been approached; the current *Standards for Educational and Psychological Testing* evolved after the first version, more than 60 years ago (American Psychological Association 1954). The current *Standards* approach validity as a unitary concept, with five sources of validity recognized (Sireci and Padilla 2014): based on internal structure, based on test content, based on the relation to other variables, based on response processes and based on the consequences of testing. Although these are not considered to present distinct types of validity, an inclusive evaluation of the instrument includes these different sources of evidence in a coherent account (American Educational Research Association et al. 2014).

Validity evidence based on the internal structure includes three basic aspects: dimensionality, measurement invariance and reliability (Rios and Wells 2014). To assess dimensionality, one can opt for several factor analytic methods; however, confirmatory factor analysis (Brown 2015) is the most comprehensive approach for comparing observed and hypothesized test structures, as it evaluates the relationships between items and the latent variables (theoretical constructs) and which items should be measured (Bollen 1989).

Measurement invariance assesses whether an instrument is fair for different subgroups from a psychometric perspective (van de Schoot et al. 2012), such as occupations (Sinval et al. 2018), countries (Reis et al. 2015), genders (Marsh et al. 2010) and other groups.

It can also be evaluated using different statistical approaches, with multigroup confirmatory factor analysis being the most popular (Davidov et al. 2014). This approach consists of setting increasingly constrained sets of structural equation models, and comparing the more restricted models with the less restricted models (van de Schoot et al. 2015).

Since the validity of scores depends on their reliability (American Educational Research Association et al. 2014), without reliability we can't have appropriate validity evidence (Kaplan and Saccuzzo 2013). It can be evaluated with different techniques, although the most usual is through internal consistency estimates, such as Cronbach's α, Revelle's β or McDonald's ω_h (Zinbarg et al. 2005). It provides evidence about the consistency of the test scores across repeated administrations (American Educational Research Association et al. 2014).

Participants

A sample of 908 Portuguese first-year university students (ages ranging from 17 to 58 years; $M = 19.41$; $SD = 4.79$; $Mdn = 18$) from a public university in the north of Portugal was used to evaluate the psychometric properties of the USEI. These students commonly took courses in three main areas: 40.18% were from technology or engineering courses; 29.52% from economics or law courses; and 30.30% from languages or humanities. Most students were women (64.58%) and only 8.57% had a part-time or full-time occupation. With respect to parents' level of education, 50.65% of mothers had a basic education level, 30.27% a high school level and 19.08% a higher education level; meanwhile, 58.99% of fathers had a basic education level, 24.64% had a secondary level and 16.38% a higher education level.

Measures and procedures

The USEI (Marôco et al. 2016) is a self-report Likert-type (1 = "never" to 5 = "always") scale with 15 items organized in three academic engagement dimensions: behavioral (BE; e.g. *I usually participate actively in group assignments*), cognitive (CE; e.g. *I like being at school*) and emotional (EE; e.g. *I try to integrate the acquired knowledge in solving new problems*). This instrument presented good evidence of reliability and factorial, convergent and discriminant validity evidence in a previous research study (Marôco et al. 2016). Exploratory and confirmatory factor analyses confirm systematically specific items for each dimension. Reliability coefficients in terms of the consistency of items are above .63 (ordinal omega values) and above .69 (ordinal alpha values) for three dimensions.

A non-probabilistic convenience sample was considered, with the inclusion criterion being students entering university. Data were collected in the classroom context with the permission and collaboration of teachers. The aims of the study were presented, and confidentiality was ensured. The participants provided informed consent stating their voluntary agreement to participate in the study. Ten minutes were enough to fill in the inventory and give some personal information for sample characterization.

Data analysis

All statistical analysis was performed with R (R Core Team 2018) and RStudio (RStudio Team 2017). The descriptive statistics were obtained using the *skimr* package (Rubia et al. 2017). Confirmatory factor analysis (CFA) was conducted to evaluate the psychometric properties of the data gathered with the USEI, namely its internal structure validity evidence. CFA was performed with the *lavaan* package (Rosseel 2012) using the

weighted least squares means and variances (WLSMV) estimation method, which is indicated for nonlinear response scales. Internal consistency reliability estimates for ordinal variables, average variance extracted (AVE) and heterotrait-monotrait (HTMT) were calculated using the *semTools* package (semTools Contributors 2016), while Mardia's Kurtosis (Mardia 1970) was assessed using the *psych* package (Revelle 2017).

The CFA was conducted to verify whether the proposed three-factor structure presented an adequate fit for the study sample data. We used as goodness-of-fit indices the TLI (Tucker-Lewis Index), $\chi^{2/}$df (ratio of chi-square to degrees of freedom), the NFI (Normed Fit Index), the CFI (Comparative Fit Index) and the RMSEA (Root Mean Square Error of Approximation). The fit of the model was considered good for CFI, NFI and TLI values above .95 and RMSEA values below .06 (Hu and Bentler 1999; Marôco 2014).

To analyze convergent validity evidence, the AVE was estimated as described in Fornell and Larcker (1981). Values of $AVE \geq .5$ were considered acceptable indicators of convergent validity evidence. To determine whether the items that are manifestations of a factor were not strongly correlated with other factors, discriminant validity evidence was assessed. Acceptable discriminant validity evidence was assumed when for two factors x and y, AVE_x and $AVE_y \geq \rho^2_{xy}$ (squared correlation between the factors x and y), or when the HTMT (Henseler et al. 2015) ratio of correlations is higher than .85 (Kline 2016).

The reliability of the internal scores evidence was assessed through internal consistency measures. The ordinal Cronbach's alpha coefficient (α) and composite reliability (CR) were calculated. Since alpha has been shown to present evidence of a measure's internal consistency only when the assumptions of the essentially tau-equivalent model are obtained (Revelle and Zinbarg 2009), the ordinal coefficient omega (ω) for each factor (Raykov 2001;

Zumbo et al. 2007) and the hierarchical omega (ω_h) coefficient (Green and Yang 2009; Kelley and Pornprasertmanit 2016; McDonald 1999) were also calculated. Higher alpha values are desirable, although excessively high values of alpha aren't recommended, as this reveals unnecessary repetition and overlap (Streiner 2003). Values of $CR \geq .7$ were considered to be satisfactory indicators of internal consistency (Marôco 2014). Omega values show evidence of how much of the overall variance of a factor in the data that is due to that specific factor, ω, was calculated for each of the three factors. As regards the ω_h, a higher value will indicate a stronger influence of the latent variable common to all of the factors, and that the observed scale scores generalize to scores for the common latent variable (Zinbarg et al. 2007). The second-order factor reliability was also calculated using the omega coefficient (Jorgensen et al. 2018).

The measurement invariance of the second-order model was assessed with the *lavaan* package (Rosseel 2012), and we established a set of comparisons within a group of seven different models based on the recommendations for ordinal variables (Millsap and Yun-Tein 2004) and for second-order models (Chen et al. 2005). An initial configural model was set, which served as a baseline (configural invariance) for further equivalence testing (Edwards et al. 2017). Next, metric invariance of the first-order factor loadings was tested with the items' loadings forced to be equal across groups; this assessed whether the subgroups attribute the same meaning to the different instrument items. The next step consisted in forcing the second-order factor loadings to be equal across groups; this checked whether the subgroups give the same meaning to the factors that compose the second-order latent factor. Afterwards, scalar invariance of the first-order factors was tested, where thresholds were added to be equal across groups (Millsap and Yun-Tein 2004). If scalar invariance was

obtained, it meant that the means or the thresholds of the items are also equal across the subgroups, enabling comparisons between the different subgroups. Next, scalar invariance of the second-order latent factor was tested, where the intercepts of the first-order latent variables were forced to be equal across groups. This checked whether the first-order latent levels were equal across groups. Usually, this was enough for measurement invariance, since the next levels are too restrictive (Marôco 2014). After, the disturbances of first-order factors were established as being equal across groups, to verify if the explained variances for the first-order latent factors were equal across groups. Finally, if residual variances were also added to be equal across groups without statistically significant differences, full uniqueness measurement invariance was obtained, which means that the explained variance for all items didn't change in regard to the subgroup (van de Schoot et al. 2012). Invariance across the different levels can be assessed using two different criteria: the ΔCFI < .01 between constrained and free models(Cheung and Rensvold 2002), and the $\Delta\chi^2$ test comparing the fit of the constrained vs. free models is not statistically significant (Satorra and Bentler 2001).

Results

Items' distributional properties

Summary measures, skewness (*sk*), kurtosis (*ku*) and a histogram for each of the 15 items are presented (Table 1) and were used to judge distributional properties and psychometric sensitivity. Absolute values of *ku* smaller than 7 and *sk* smaller than 3 were considered an indication of not strong deviations from the normal distribution (Finney and DiStefano 2013). Mardia's multivariate kurtosis for the 15 items of the USEI was 37.5; $p < .001$. All possible answer values for each item are also present, and no outliers were deleted.

Also interesting is the reduced number of missing answers from the 15 items (11 omissions from item 10 "*My classroom is an interesting place to be*").

The items' distributional coefficients are indicative of appropriate psychometric sensitivity, as it would be expected that these items would follow an approximately normal distribution in the population under study. Despite these univariate and multivariate normality indicators, the WLSMV estimator was used to account for the ordinal level of measurement of the items, which can be done without concerns about this estimate.

Insert Table 1

Factorial validity evidence

In light of the previous researches on the USEI structure confirming the existence of three dimensions, a confirmatory factor analysis was performed. The hypothesized three-factor model's fit with the data was good (Figure 1; correlations between latent variables, and factor loadings for each item are shown), since CFI, NFI and TLI values were greater than .95, and RMSEA values were less than .06. It is also important that the factor loadings of all items are greater than .50, except for item 6 (the only reversed coded item in the instrument).

Figure 1 – Confirmatory factor analysis of the University Students Engagement Inventory (15 items) with first-year Portuguese university students ($\chi^2(87) = 286.665$, $p < .001$, $n = 871$, $CFI = .987$, $TLI = .985$, $NFI = .982$, $RMSEA = .051$, $P(\text{RMSEA} \leq .05) = .356$, IC90].045; .058[. R – Reversed.

Insert Figure 1 here

Convergent validity evidence

The average variance extracted (AVE) was acceptable for EE (.54), nearly acceptable for CE (.49) and low for BE (.31). The convergent validity evidence was acceptable for the CE and EE factors and unsatisfactory for the BE factor.

Discriminant validity evidence

Comparing data from these three dimensions, the AVE for EE (AVE_{EE} = .54) was greater than $r^2_{BE.EE}$ (.36), but the AVE_{BE} = .31 was lower, the AVE_{CE} = .49 and AVE_{EE} = .54 were both greater than $r^2_{EE.CE}$ = .24, and the AVE_{BE} = .31 and AVE_{CE} = .49 were both less than $r^2_{BE.CE}$ = .52. The discriminant validity evidence was good for CE and EE, insufficient for BE and EE, and poor for BE and CE. With regard to the HTMT criterion, the $HTMT_{BE.EE}$ = .60, $HTMT_{BE.CE}$ = .73 and $HTMT_{EE.CE}$ = .51, with all being below the recommended threshold. Together, these findings detect strong correlations/overlap among the three latent constructs. This points to a possible higher-order latent factor.

Second-order construct

We tested the possible existence of a higher-order latent variable, the meta-construct *academic engagement*, which was hypothesized by the original authors (Marôco et al. 2016), and suggested also by our lack of discriminant validity evidence findings. In regard to the USEI with a second-order latent factor, overall the goodness-of-fit indices were good (Figure 2; gamma between the second-order latent factor and the first-order latent factors, and factor loadings for each item are shown). The structural weights for the *academic engagement*

second-order factor model were medium/high: behavioral engagement ($\gamma = 0.93$; $p < .001$); emotional engagement ($\gamma = 0.64$; $p < .001$); and cognitive engagement ($\gamma = 0.77$; $p < .001$).

Figure 2 – Confirmatory factor analysis of the University Students Engagement Inventory (second-order model – 15 items) with first-year Portuguese university students ($\chi^2(87) = 286.665$, $p < .001$, $n = 871$, *CFI* = .987, *TLI* = .985, *NFI* = .982, *RMSEA* = .051, *P*(RMSEA ≤ .05) = .356, IC90].045; .058[. R – Reversed.

Insert Figure 2 here

Reliability: Internal consistency evidence

In terms of the hypothesized reliability evidence, the results suggest good evidence of internal consistency reliability (Table 2). The alpha values were higher than the omega values for all factors and for the total scale. The hierarchical omega for the total scale was good ($\omega_h = .85$), which suggests a well-defined latent variable, thereby evidencing that this latent variable is more likely to be stable across studies, which also suggests that the general factor *academic engagement* is the dominant source of systematic variance (Rodriguez et al. 2016).

Insert Table 2

The internal consistency reliability of the second-order latent variable was good. The proportion of observed variance explained by the second-order factor after controlling for the uniqueness of the first-order factor ($\omega_{partial\ L1}$) was .87, the proportion of the second-order factor explaining the variance of the first-order factor level (ω_{L2}) was .87 and the proportion of the second-order factor explaining the total score (ω_{L1}) was .72.

Measurement invariance by gender and scientific area of graduation

Finally, to detect whether the same second-order latent model holds in different scientific areas of graduation and genders, a group of nested models with indications of equivalence is needed. The hypothesized full-scale invariance was supported for gender (Table 3) using the Cheung and Rensvold (2002) ΔCFI criterion, while the $\Delta\chi^2$ criterion supported only the second-order metric invariance. In regard to the hypothesized structural invariance among different areas of study, full-scale invariance was supported by the ΔCFI criterion, nevertheless the ΔCFI value for the first comparison was marginal at .010, although the $\Delta\chi^2$ supported it, after we continued with the comparisons; the $\Delta\chi^2$ criterion allowed only the first-order metric invariance (see Table 4). In both cases, the ΔCFI criterion was preferable, since the $\Delta\chi^2$ is too restrictive (Marôco 2014).

Insert Table 3

Insert Table 4

Discussion

Hypotheses findings

This study obtained findings that allow our H1 to be confirmed, since the data gathered with the USEI presented good psychometric properties in terms of validity evidence based on the internal structure, something that was observed in other studies with this instrument, which obtained acceptable/good overall goodness of fit (Costa et al. 2014) and

good overall goodness of fit (Marôco et al. 2016). The confirmatory factor analysis presented good evidence about factorial validity, since goodness-of-fit indices values ranged from very good to good, and only item 6 had a lambda of less than .50. Analyzing its content, item 6 is the only one reverse coded, which suggests that it should be presented in the same direction as the other items in the future. Marôco, Marôco and Campos (2014) report this kind of improvement in the items' correlations in student burnout (an opposite construct to academic engagement). The USEI's convergent validity evidence is acceptable and the AVE values were good for the EE dimension, marginally acceptable for CE and less than acceptable for BE. These values show that the items of each dimension were good manifestations of the factors they load onto. The discriminant validity evidence of the instrument was acceptable for two of the three factors. The lack of discriminant validity evidence for BE may be due to our sample being composed only of freshmen; in the original USEI study (Marôco et al. 2016) with students from other academic years, this lack of discriminant validity evidence was not observed.

Our H2 was confirmed, something that has been tested by the original authors, with whom our results were aligned in terms of structural weights, with behavioral engagement having the highest gamma, followed by cognitive engagement and finally emotional engagement (Marôco et al. 2016).

With regard to H3 and H4, our results brought evidence that allows comparisons to be established between male and female genders using the USEI, and between first-year students from technology or engineering courses, from economics or law courses, and from languages or humanities. This finding was a novelty of our study, and is useful since previous studies only assessed engineering students (Costa et al. 2014; Costa and Marôco 2017) or – even with a sample from different courses – didn't test measurement invariance for the

scientific area of graduation (Marôco et al. 2016). Another novelty of this study was the test of the second-order measurement invariance, since the only study that has tested measurement invariance using this instrument (Marôco et al. 2016) did so only to compare the structure between two independent samples without comparing specific scientific areas of graduation, and regarding the first-order model. This finding will enable future comparisons among these different groups to verify possible differences and their impact on academic adjustment and achievement.

With regard to the evidence obtained about reliability, it was good for CR, ordinal α, ordinal ω and ω_h, suggesting adequate reliability of the data measured with the USEI. Our results confirm our H5, and – nevertheless – are aligned with what was found in other studies, where BE obtained lower reliability estimates than EE and CE (Costa et al. 2014; Marôco et al. 2016).

Academic engagement is a relevant construct for describing student adaptation and achievement in higher education. Engaged students tend to invest more in their performance and develop strategies to persist in and to self-regulate their learning (Christenson and Reschly 2010; Dılekmen 2007; Fredricks et al. 2011; Klem and Connell 2004). Consequently, better academic success is expected (Lee 2014). In the literature, some consensus exists defining academic engagement as a multidimensional construct, integrating behavioral, emotional and cognitive dimensions (Fredricks et al. 2004). Our data from the USEI confirm these three dimensions for describing students' academic engagement. Albite the second-order construct (academic engagement) presents higher path loadings in the behavioral dimension ($\gamma = .93$) than emotional and cognitive engagement, respectively $\gamma = .64$ and $\gamma = .77$. The differences are pertinent and are in line with the expected, since our sample was constituted by freshmen. The literature suggests that the first-years students have less

maturity and autonomy to cope with the challenges of higher education (Bernardo et al. 2017; Pascarella and Terenzini 2005). The first-year college students have their academic engagement more expressed on behavioral terms, which can be seen on the academic routines and tasks (e.g. attend to classes, group assignments).

Based on the foregoing discussion, we conclude that the USEI presents good validity evidence about its internal structure, presenting promising results for future studies related to other sources of validity and different university students' samples. This instrument can become an interesting tool for education and psychology researchers for analyzing the relationship between the different types of academic engagement and other personal and academic variables important for students' adjustment and academic achievement.

Although domain-specific subject areas aren't included in the instrument, they may contribute to understanding the extent to which engagement is content-specific, and to what extent it represents a general engagement tendency (Fredricks et al. 2004). Since this was a study carried out with a sample of university students from different courses, it wasn't desirable to have a different version for each course, and it wasn't practical either because of time and resource constraints. If one wants to understand and study a specific academic engagement dimension, this kind of more inclusive instrument might be insufficient, although if the goal of the study is to obtain a single measure for each of the three types of academic engagement, this instrument may be a good choice, since it addresses each construct with few items, and the last word is given to the researcher.

Theoretical Implications

This study presents some theoretical findings that can enable a better understanding of academic engagement as a multidimensional construct. USEI revealed a three-factor

structure that appears to be indicative of a higher-order construct, academic engagement. This makes USEI unique regarding the potential of its conceptualization of academic engagement as a meta-construct (Fredricks and McColskey 2012b), which is important to define well in terms of its subdimensions (Fredricks et al. 2016). The results emphasize that this conception of academic engagement works well in different scientific areas of college graduation courses. There are some other subject specific instruments (Kong et al. 2003; Wigfield et al. 2008), but USEI has the particularity of being a general measure of academic engagement for university students.

This is the first report addressing the USEI validity for students majoring different study areas. The behavioral and emotional components of academic engagement in this instrument didn't present the desired discriminant validity evidence, appearing to be somehow related to their content. Our validity evidence supports a consistent alignment with the academic engagement construct definition, showing good psychometric properties for the study sample. As a convergence or product of motivation and active learning behaviors, academic engagement works as a relevant variable with a strong impact in predicting the student's permanence and success in completing his or her course in higher education (Alrashidi et al. 2016; Barkley 2010; Kuh 2001).

Practical Implications

As for practical implications, USEI can be considered a tool with good psychometric properties that can measure the perceptions of academic engagement behaviors, emotions and internal cognitions in first-year university students. It is an instrument that was specifically designed for university students and it is available for free. This can be done across groups of different scientific areas of graduation and genders, without losing the

desirable measurement invariance that enables direct comparisons of scores between those groups. This is something that hasn't been done before across gender or the scientific area of graduation. These findings together can bring confidence to the measures obtained using the USEI, knowing the academic engagement predictive relation with other variables. For example, Costa and Marôco (2017) found that the emotional subdimension of academic engagement had a statistically significant relation with students' dropout thoughts. Consequently, this is an important implication, since USEI can be useful to assess interventions for specific dimensions of students' engagement. USEI is particularly useful for measuring cognitive engagement and emotional engagement, that are not directly observed (Fredricks and McColskey 2012b). With USEI these subdimensions don't need to be inferred from behavioral indicators or teacher rating scales, avoiding potential inferences through those other methods (Appleton et al. 2006).

Self-report instruments have several advantages over other methods, they are practical and relatively low cost tools for group or large-scale assessments (Mandernach 2015). This allows to obtain data over several waves and establish different types of comparisons (e.g. universities, courses). The large-scale assessment of academic engagement enables teachers, policymakers and administrative boards to assess students' learning status and their academic life experiences (Coates 2005), making it possible to obtain relevant instructional feedback to the institution's decision-makers and to the students themselves regarding the measured constructs (Banta et al. 2009; Kember and Leung 2009). In this sense, due to its psychological, contextual nature and complexity, academic engagement assessment should take a multidimensional approach considering the behavioral, emotional and cognitive aspects (Alrashidi et al. 2016; Mandernach 2015). This multidimensional approach allows for differential analysis. For example, on the levels and types of investment in relation

to the scientific areas and to differentiated subgroups of students according to their socio-cultural origin or their vocational career projects. With the Bologna Declaration (1999), governments in European countries advocated for higher education to value and be based on the active participation of students in their skills development and learning. This perspective benefits from brief and multidimensional instruments that ensure a large-scale assessment of the students' levels of academic engagement as related to their behavioral, cognitive and emotional aspects.

Conclusions

All the research hypotheses were confirmed, pointing to the validity evidence of the obtained findings, something that goes in line with other previous studies (Costa et al. 2014; Costa and Marôco 2017; Marôco et al. 2016). There seems to be evidence that the USEI is an appropriate psychometric instrument for the academic engagement framework adopted, which is multidimensional and comprised of observable behaviors, emotions and internal cognitions. Thus, it can help to capitalize on the potential of academic engagement as a multidimensional construct (Fredricks 2015), with a higher-order dimension, academic engagement. Our findings bring clarity regarding the psychometric properties of this promising instrument, which can successfully measure the three different kinds of academic engagement from the most adopted theoretical framework. This is the first instrument that enables Portuguese university students to do so. Due to its reduced number of items this instrument can be adequate to research proposes in large scale related with academic engagement, and to practical purposes at the intervention levels can allow to identify dimensions where teachers and university staff can design interventions based on the specificities of each scientific areas or students' subgroups.

Future studies should address longitudinal research designs, such as longitudinal measurement invariance and measurement invariance for public/private universities, with students from different graduation years, something that isn't often implemented as it should be, since it is a condition for making proper comparisons between different groups (Davidov et al. 2014). Future studies should also look at transcultural validity of the USEI in different languages, other than the European/Brazilian Portuguese for which the USEI was initially developed.

Also, other kinds of validity evidence should be addressed, such as evidence of validity based on relationships with measures of other variables like student achievement, drop-out, burnout and well-being (McCoach et al. 2013). Our sample only included first-year students from a Portuguese public university, and it is desirable that other and more diverse scientific areas of graduation should be included, such as students from private universities and different grades, and also students with another status (such as a student worker).

Funding This work received no grants or assistance from a funding body.

Compliance with Ethical Standards

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the Institutional Ethics Research Committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

548 **Informed Consent** Informed consent was obtained from all individual participants
549 included in the study.
550

References

Alrashidi, O., Phan, H. P., & Ngu, B. H. (2016). Academic engagement: An overview of its definitions, dimensions, and major conceptualisations. *International Education Studies*, *9*(12), 41–52. doi:10.5539/ies.v9n12p41

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington, DC, USA: American Psychological Association.

American Psychological Association. (1954). Technical recommendations for psychological tests and diagnostic techniques. *Psychological Bulletin*, *51*(2), 201–238. doi:10.1037/h0053479

Antúnez, Á., Cervero, A., Solano, P., Bernardo, I., & Carbajal, R. (2017). Engagement: A new perspective for reducing dropout through self-regulation. In J. A. González-Pienda, A. Bernardo, J. C. Núñez, & C. Rodríguez (Eds.), *Factors affecting academic performance* (pp. 25–46). New York, NY, USA: Nova Science Publishers.

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. doi:10.1002/pits.20303

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. doi:10.1016/j.jsp.2006.04.002

Bakker, A. B., Schaufeli, W. B., Leiter, M. P., & Taris, T. W. (2008). Work engagement: An emerging concept in occupational health psychology. *Work & Stress*, *22*(3), 187–200. doi:10.1080/02678370802393649

Banta, T. W., Pike, G. R., & Hansen, M. J. (2009). The use of engagement data in accreditation, planning, and assessment. *New Directions for Institutional Research*, *2009*(141), 21–34. doi:10.1002/ir.284

Barkley, E. F. (2010). *Student engagement techniques: A handbook for college faculty*. San Francisco, CA, USA: John Wiley & Sons.

Bernardo, A., Cervero, A., Esteban, M., Tuero, E., Casanova, J. R., & Almeida, L. S. (2017). Freshmen program withdrawal: Types and recommendations. *Frontiers in Psychology*, *8*, 1–11. doi:10.3389/fpsyg.2017.01544

Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY, USA: Wiley. doi:10.1002/9781118619179

Bologna Declaration. The European higher education area. Joint Declaration of the European Ministers of Education, June 19, Bologna. (1999). https://www.eurashe.eu/library/bologna_1999_bologna-declaration-pdf/

Bresó, E., Schaufeli, W. B., & Salanova, M. (2011). Can a self-efficacy-based intervention decrease burnout, increase engagement, and enhance performance? A quasi-experimental study. *Higher Education*, *61*(4), 339–355. doi:10.1007/s10734-010-9334-6

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY, USA: The Guildford Press.

Carter, C. P., Reschly, A. L., Lovelace, M. D., Appleton, J. J., & Thompson, D. (2012). Measuring student engagement among elementary students: Pilot of the Student Engagement Instrument— Elementary version. *School Psychology Quarterly*, *27*(2),

61–73. doi:10.1037/a0029229
Chapman, C., Laird, J., Ifill, N., & KewalRamani, A. (2011). *Trends in high school dropout and completion rates in the United States: 1972 – 2009* (No. compendium report 2012-006). National Center for Education Statistics. http://nces.ed.gov/pubsearch
Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*(3), 471–492. doi:10.1207/s15328007sem1203_7
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. doi:10.1207/S15328007SEM0902_5
Christensen, M. (2017). Work engagement and job crafting. In M. Christensen, P. Ø. Saksvik, & M. Karanika-Murray (Eds.), *The positive side of occupational health psychology* (pp. 79–90). Cham, Switzerland: Springer. doi:10.1007/978-3-319-66781-2_7
Christenson, S. L., & Reschly, A. L. (2010). Check & connect: Enhancing school completion through student engagement. In B. Doll, W. Pfohl, & J. Yoon (Eds.), *Handbook of youth prevention science* (pp. 327–334). New York, NY, USA: Routledge.
Christenson, S. L., Reschly, A. L., & Wylie, C. (Eds.). (2012). *Handbook of research on student engagement*. New York, NY, USA: Springer. doi:10.1007/978-1-4614-2018-7
Coates, H. (2005). The value of student engagement for higher education quality assurance. *Quality in Higher Education*, *11*(1), 25–36. doi:10.1080/13538320500074915
Coetzee, M., & Oosthuizen, R. M. (2012). Students' sense of coherence, study engagement and self-efficacy in relation to their study and employability satisfaction. *Journal of Psychology in Africa*, *22*(3), 315–322. doi:10.1080/14330237.2012.10820536
Costa, A. R., Araújo, A. M., & Almeida, L. S. (2014). Envolvimento académico de estudantes de engenharia: Contributos para a validação interna e externa de uma escala de avaliação. *Revista Eletrónica de Psicologia, Educação e Saúde*, *4*(1), 142–155. http://www.revistaepsi.com
Costa, A. R., & Marôco, J. (2017). Inventário de envolvimento académico dos estudantes do Ensino Superior (USEI). In L. S. Almeida, M. R. Simões, & M. M. Gonçalves (Eds.), *Adaptação, desenvolvimento e sucesso académico dos estudantes do ensino superior: Instrumentos de avaliação* (pp. 33–44). Braga, Portugal: ADIPSIEDUC.
Crutzen, R., & Peters, G.-J. Y. (2017). Scale quality: Alpha is an inadequate estimate and factor-analytic evidence is needed first of all. *Health Psychology Review*, *11*(3), 242–247. doi:10.1080/17437199.2015.1124240
Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, *40*(1), 55–75. doi:10.1146/annurev-soc-071913-043137
Dılekmen, M. (2007). Orientation program and adaptation of university students. *Psychological Reports*, *101*(7), 1141–1144. doi:10.2466/PR0.101.7.1141-1144
Eccles, J., & Wang, M.-T. (2012). So what is student engagement anyway? In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 133–145). Boston, MA, USA: Springer. doi:10.1007/978-1-4614-2018-7_6
Edwards, M. C., Houts, C. R., & Wirth, R. J. (2017). Measurement invariance, the lack thereof, and modeling change. *Quality of Life Research*, 1–9. doi:10.1007/s11136-

017-1673-7
Elmore, G. M., & Huebner, E. S. (2010). Adolescents' satisfaction with school experiences: Relationships with demographics, attachment relationships, and school engagement behavior. *Psychology in the Schools*, *47*(6), 525–537. doi:10.1002/pits.20488
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). Boston, MA, USA: Springer. doi:10.1007/978-1-4614-2018-7_5
Finney, S. J., & DiStefano, C. (2013). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 439–492). Charlotte, NC, USA: Information Age Publishing.
Fredricks, J. A. (2015). Academic engagement. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., Vol. 1, pp. 31–36). San Diego, CA, USA: Elsevier. doi:10.1016/B978-0-08-097086-8.26085-6
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. doi:10.3102/00346543074001059
Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. *Learning and Instruction*, *43*, 1–4. doi:10.1016/j.learninstruc.2016.02.002
Fredricks, J. A., & McColskey, W. (2012a). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 763–782). Boston, MA, USA: Springer. doi:10.1007/978-1-4614-2018-7_37
Fredricks, J. A., & McColskey, W. (2012b). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 763–782). Boston, MA, USA: Springer. doi:10.1007/978-1-4614-2018-7_37
Fredricks, J. A., McColskey, W., Meli, J., Mordica, J., Montrosse, B., & Mooney, K. (2011). *Measuring student engagement in upper elementary through high school: A description of 21 instruments* (No. Issues & Answers report, REL 2011–No. 098). Washington, DC, USA: Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. http://ies.ed.gov/ncee/edlabs
Gilardi, S., & Guglielmetti, C. (2011). University life of non-traditional students: Engagement styles and impact on attrition engagement styles and impact on attrition. *The Journal of Higher Education*, *82*(1), 33–53. doi:10.1080/00221546.2011.11779084
Green, S. B., & Yang, Y. (2009). Reliability of summed item scores using structural equation modeling: An alternative to coefficient alpha. *Psychometrika*, *74*(1), 155–167. doi:10.1007/s11336-008-9099-3
Handelsman, M. M., Briggs, W. L., Sullivan, N., & Towler, A. (2005). A measure of college student course engagement. *The Journal of Educational Research*, *98*(3), 184–192. doi:10.3200/JOER.98.3.184-192

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. doi:10.1007/s11747-014-0403-8

Hirschi, A. (2012). Callings and work engagement: Moderated mediation model of work meaningfulness, occupational identity, and occupational self-efficacy. *Journal of Counseling Psychology*, *59*(3), 479–485. doi:10.1037/a0028949

Hodson, P., & Thomas, H. (2003). Quality assurance in Higher Education: Fit for the new millennium or simply year 2000 compliant? *Higher Education*, *45*(3), 375–387. doi:10.1023/A:1022665818216

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. doi:10.1080/10705519909540118

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2018). semTools: Useful tools for structural equation modeling (R package version 0.4-15.930) [Computer software]. https://cran.r-project.org/package=semTools

Kahu, E. R. (2013). Framing student engagement in higher education. *Studies in Higher Education*, *38*(5), 758–773. doi:10.1080/03075079.2011.598505

Kaplan, R. M., & Saccuzzo, D. P. (2013). *Psychological testing: Principles, applications & issues* (8th ed.). Belmont, CA, USA: Wadsworth.

Kelley, K., & Pornprasertmanit, S. (2016). Confidence intervals for population reliability coefficients: Evaluation of methods, recommendations, and software for composite measures. *Psychological Methods*, *21*(1), 69–92. doi:10.1037/a0040086

Kember, D., & Leung, D. Y. P. (2009). Development of a questionnaire for assessing students' perceptions of the teaching and learning environment and its use in quality assurance. *Learning Environments Research*, *12*(1), 15–29. doi:10.1007/s10984-008-9050-7

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health*, *74*(7), 262–273. doi:10.1111/j.1746-1561.2004.tb08283.x

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY, USA: The Guildford Press.

Kong, Q., Wong, N.-Y., & Lam, C.-C. (2003). Student engagement in mathematics: Development of instrument and validation of construct. *Mathematics Education Research Journal*, *15*(1), 4–21. doi:10.1007/BF03217366

Krause, K.-L., & Coates, H. (2008). Students' engagement in first-year university. *Assessment & Evaluation in Higher Education*, *33*(5), 493–505. doi:10.1080/02602930701698892

Kuh, G. D. (2001). Assessing what really matters to student learning: Inside The National Survey of Student Engagement. *Change: The Magazine of Higher Learning*, *33*(3), 10–17. doi:10.1080/00091380109601795

Kuh, G. D. (2005). Student engagement in the first year of college. In M. L. Upcraft, J. N. Gardner, & B. O. Barefoot (Eds.), *Challenging and supporting the first-year student: A handbook for improving the first year of college* (pp. 86–107). San Francisco, CA, USA: Jossey-Bass. http://www.ncsu.edu/uap/transition_taskforce/documents/documents/ReferenceMaterialPart3.pdf

Kuh, G. D. (2009). The National Survey of Student Engagement: Conceptual and empirical

738 foundations. *New Directions for Institutional Research*, *2009*(141), 5–20.
739 doi:10.1002/ir.283
740 Lee, J.-S. (2014). The relationship between student engagement and academic performance:
741 Is it a myth or reality? *The Journal of Educational Research*, *107*(3), 177–185.
742 doi:10.1080/00220671.2013.807491
743 Lin, S.-H., & Huang, Y.-C. (2018). Assessing college student engagement: Development
744 and validation of the student course engagement scale. *Journal of Psychoeducational*
745 *Assessment*, *36*(7), 694–708. doi:10.1177/0734282917697618
746 Mandernach, B. J. (2015). Assessment of student engagement in higher education: A
747 synthesis of literature and assessment tools. *International Journal of Learning,*
748 *Teaching and Educational Research*, *12*(2), 1–14. doi:10.1073/pnas.94.17.9499
749 Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications.
750 *Biometrika*, *57*(3), 519–530. doi:10.1093/biomet/57.3.519
751 Marôco, J. (2014). *Análise de equações estruturais: Fundamentos teóricos, software &*
752 *aplicações* (2nd ed.). Pêro Pinheiro, Portugal: ReportNumber.
753 Maroco, J., Maroco, A. L., & Campos, J. A. D. B. (2014). Student's Academic Efficacy or
754 Inefficacy? An Example on How to Evaluate the Psychometric Properties of a
755 Measuring Instrument and Evaluate the Effects of Item Wording. *Open Journal of*
756 *Statistics*, *04*(06), 484–493. doi:10.4236/ojs.2014.46046
757 Marôco, J., Marôco, A. L., Campos, J. A. D. B., & Fredricks, J. A. (2016). University
758 student's engagement: Development of the University Student Engagement Inventory
759 (USEI). *Psicologia: Reflexão e Crítica*, *29*(21), 1–12. doi:10.1186/s41155-016-0042-8
760 Marsh, H. W., Lüdtke, O., Muthén, B. O., Asparouhov, T., Morin, A. J. S., Trautwein, U.,
761 & Nagengast, B. (2010). A new look at the big five factor structure through
762 exploratory structural equation modeling. *Psychological assessment*, *22*(3), 471–491.
763 doi:10.1037/a0019227
764 McCoach, D. B., Gable, R. K., & Madura, J. P. (Eds.). (2013). Evidence based on relations
765 to other variables: Bolstering the empirical validity arguments for constructs. In
766 *Instrument development in the affective domain* (3rd ed., pp. 209–248). New York,
767 NY, USA: Springer. doi:10.1007/978-1-4614-7135-6_6
768 McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ, USA: Routledge.
769 doi:10.4324/9781410601087
770 Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical
771 measures. *Multivariate Behavioral Research*, *39*(3), 479–515.
772 doi:10.1207/S15327906MBR3903_4
773 National Research Council, & Institute of Medicine. (2004). *Engaging schools: Fostering*
774 *school students' motivation to learn*. Washington, DC, USA: National Academies
775 Press. doi:10.17226/10421
776 Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of*
777 *research* (Vol. 2). San Francisco, CA, USA: Jossey-Bass.
778 R Core Team. (2018). R: A language and environment for statistical computing (version
779 3.5.1) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing.
780 https://www.r-project.org/
781 Raykov, T. (2001). Estimation of congeneric scale reliability using covariance structure
782 analysis with nonlinear constraints. *The British Journal of Mathematical and*
783 *Statistical Psychology*, *54*, 315–323. doi:10.1348/000711001159582
784 Reis, D., Xanthopoulou, D., & Tsaousis, I. (2015). Measuring job and academic burnout

785 with the Oldenburg Burnout Inventory (OLBI): Factorial invariance across samples
786 and countries. *Burnout Research*, *2*(1), 8–18. doi:10.1016/j.burn.2014.11.001
787 Reschly, A. L., & Christenson, S. L. (2012a). Jingle, jangle, and conceptual haziness:
788 Evolution and future directions of the engagement construct. In S. L. Christenson, A.
789 L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 3–
790 19). Boston, MA, USA: Springer. doi:10.1007/978-1-4614-2018-7_1
791 Reschly, A. L., & Christenson, S. L. (2012b). Jingle, jangle, and conceptual haziness:
792 Evolution and future directions of the engagement construct. In S. L. Christenson, A.
793 L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 3–
794 19). Boston, MA, USA: Springer. doi:10.1007/978-1-4614-2018-7_1
795 Revelle, W. (2017). psych: Procedures for psychological, psychometric, and personality
796 research (R package version 1.7.8) [Computer software]. Evanston, IL: Northwestern
797 University. https://cran.r-project.org/package=psych
798 Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb:
799 Comments on sijtsma. *Psychometrika*, *74*(1), 145–154. doi:10.1007/s11336-008-9102-
800 z
801 Rios, J., & Wells, C. (2014). Validity evidence based on internal structure. *Psicothema*,
802 *26*(1), 108–16. doi:10.7334/psicothema2013.260
803 Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices
804 in the evaluation of psychological measures. *Journal of Personality Assessment*, *98*(3),
805 223–237. doi:10.1080/00223891.2015.1089249
806 Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of*
807 *Statistical Software*, *48*(2), 1–21. http://www.jstatsoft.org/v48/i02/paper
808 RStudio Team. (2017). RStudio: Integrated development for R (version 1.1.345) [Computer
809 software]. Boston, MA, USA: RStudio, Inc. http://www.rstudio.com/
810 Rubia, E. A. de la, Zhu, H., Ellis, S., Waring, E., & Quinn, M. (2017). skimr: skimr (R
811 package version 0.9000) [Computer software].
812 Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for
813 moment structure analysis. *Psychometrika*, *66*(4), 507–514. doi:10.1007/BF02296192
814 Schaufeli, W. B., & Bakker, A. B. (2010). Defining and measuring work engagement:
815 Bringing clarity to the concept. In A. B. Bakker & M. P. Leiter (Eds.), *Work*
816 *engagement: A handbook of essential theory and research* (pp. 10–24). New York,
817 NY: Psychology Press. http://psycnet.apa.org/psycinfo/2010-06187-002
818 Schaufeli, W. B., Martinez, I. M., Marques-Pinto, A., Salanova, M., & Bakker, A. B.
819 (2002). Burnout and engagement in university students: A cross-national study.
820 *Journal of Cross-Cultural Psychology*, *33*(5), 464–481.
821 doi:10.1177/0022022102033005003
822 Schaufeli, W. B., Martinez, I. M., Marques Pinto, A., Salanova, M., & Bakker, A. B.
823 (2002). Burnout and engagement in university students: A cross-national study.
824 *Journal of Cross-Cultural Psychology*, *33*(5), 464–481.
825 doi:10.1177/0022022102033005003
826 semTools Contributors. (2016). semTools: Useful tools for structural equation modeling (R
827 package version 0.4-14) [Computer software]. https://cran.r-
828 project.org/package=semTools
829 Senior, C., & Howard, C. (2015). Editorial: The state of the art in student engagement.
830 *Frontiers in Psychology*, *6*, 1–2. doi:10.3389/fpsyg.2015.00355
831 Sheppard, S. L. (2011). School engagement: A ‘danse macabre’? *Journal of Philosophy of*

Education, *45*(1), 111–123. doi:10.1111/j.1467-9752.2010.00782.x

Sinval, J., Marques-Pinto, A., Queirós, C., & Marôco, J. (2018). Work engagement among rescue workers: Psychometric properties of the Portuguese UWES. *Frontiers in Psychology*, *8*. doi:10.3389/fpsyg.2017.02229

Sireci, S. G., & Padilla, J.-L. (2014). Validating assessments: Introduction to the special section. *Psicothema*, *26*(1), 97–99. doi:10.7334/psicothema2013.255

Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, *80*(1), 99–103. doi:10.1207/S15327752JPA8001_18

van de Schoot, R., Lugtig, P., & Hox, J. J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486–492. doi:10.1080/17405629.2012.686740

van de Schoot, R., Schmidt, P., & De Beuckelaer, A. (Eds.). (2015). *Measurement invariance*. Lausanne, Switzerland: Frontiers Media. doi:10.3389/978-2-88919-650-0

Vasalampi, K., Salmela-Aro, K., & Nurmi, J.-E. (2009). Adolescents' self-concordance, school engagement, and burnout predict their educational trajectories. *European Psychologist*, *14*(4), 332–341. doi:10.1027/1016-9040.14.4.332

Wierstra, R. F. A., Kanselaar, G., van der Linden, J. L., Lodewijks, H. G. L. C., & Vermunt, J. D. (2003). The impact of the university context on European students' learning approaches and learning environment preferences. *Higher Education*, *45*(4), 503–523. doi:10.1023/A:1023981025796

Wigfield, A., Guthrie, J. T., Perencevich, K. C., Taboada, A., Klauda, S. L., McRae, A., & Barbosa, P. (2008). Role of reading engagement in mediating effects of reading comprehension instruction on reading outcomes. *Psychology in the Schools*, *45*(5), 432–445. doi:10.1002/pits.20307

Wolf-Wendel, L., Ward, K., & Kinzie, J. (2009). A tangled web of terms: The overlap and unique contribution of involvement, engagement, and integration to understanding college student success. *Journal of College Student Development*, *50*(4), 407–428. doi:10.1353/csd.0.0077

Wonglorsaichon, B., Wongwanich, S., & Wiratchai, N. (2014). The influence of students school engagement on learning achievement: A structural equation modeling analysis. *Procedia - Social and Behavioral Sciences*, *116*, 1748–1755. doi:10.1016/j.sbspro.2014.01.467

Zinbarg, R. E., Revelle, W., & Yovel, I. (2007). Estimating ωh for structures containing two group factors: Perils and prospects. *Applied Psychological Measurement*, *31*(2), 135–157. doi:10.1177/0146621606291558

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and Mcdonald's ωH: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*(1), 123–133. doi:10.1007/s11336-003-0974-7

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients Alpha and Theta for likert rating scales. *Journal of Modern Applied Statistical Methods*, *6*(1), 21–29. http://digitalcommons.wayne.edu/jmasm%5Cnhttp://digitalcommons.wayne.edu/jmasm/vol6/iss1/4

878 Table 1. Distributional properties of USEI's items

item	missing	*N*	*M*	*SD*	min	max	histogram
eng1	0	908	3.96	0.62	1	5	
eng2	2	908	4.62	0.57	1	5	
eng3	2	908	4.20	0.79	1	5	
eng4	1	908	3.35	0.91	1	5	
eng5	7	908	4.36	0.76	1	5	
eng6[R]	5	908	3.88	1.01	1	5	
eng7	1	908	3.81	0.75	1	5	
eng8	3	908	4.14	0.74	1	5	
eng9	4	908	3.91	0.79	1	5	
eng10	11	908	3.69	0.77	1	5	
eng11	2	908	3.84	0.90	1	5	
eng12	1	908	3.57	0.92	1	5	
eng13	2	908	4.07	0.84	1	5	
eng14	1	908	3.93	0.71	1	5	
eng15	2	908	4.07	0.71	1	5	

Note. R – reversed.

879

880

881

882 Table 2. Internal consistency of USEI dimensions for the Total Sample

USEI's dimensions	$\alpha_{ordinal}$	$\omega_{ordinal}$	*CR*
BE	.69	.63	.70
EE	.83	.78	.85
CE	.80	.75	.82
Total	.87	.85	-

883

Table 3. USEI's models comparisons for gender

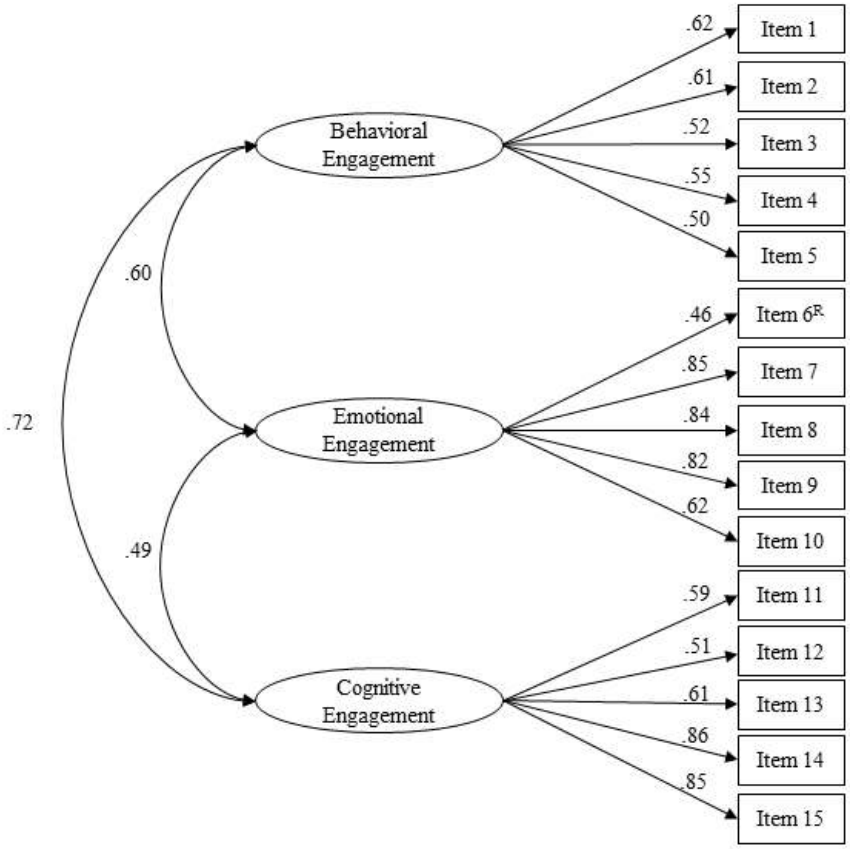
Model invariance	$\chi 2$	df	$\chi 2$/df	*CFI* scaled	$\Delta\chi 2$	Δ*CFI* scaled
Configural (factor structure)	336.29	174	1.93	.968	-	-
First-order loadings invariance	353.66	186	1.90	.969	6.94^{ns}	.001
Second-order loadings invariance	353.88	188	1.88	.970	0.12^{ns}	.001
Thresholds of measured variables	460.47	230	2.00	.964	57.61^{***}	.006
Intercepts of first-order factors invariance	499.33	233	2.14	.963	5.44^{*}	.001
Disturbances of first-order factors invariance	502.59	235	2.14	.963	1.40^{ns}	.000
Residual variances of observed variables invariance	578.33	250	2.31	.959	37.90^{***}	.004

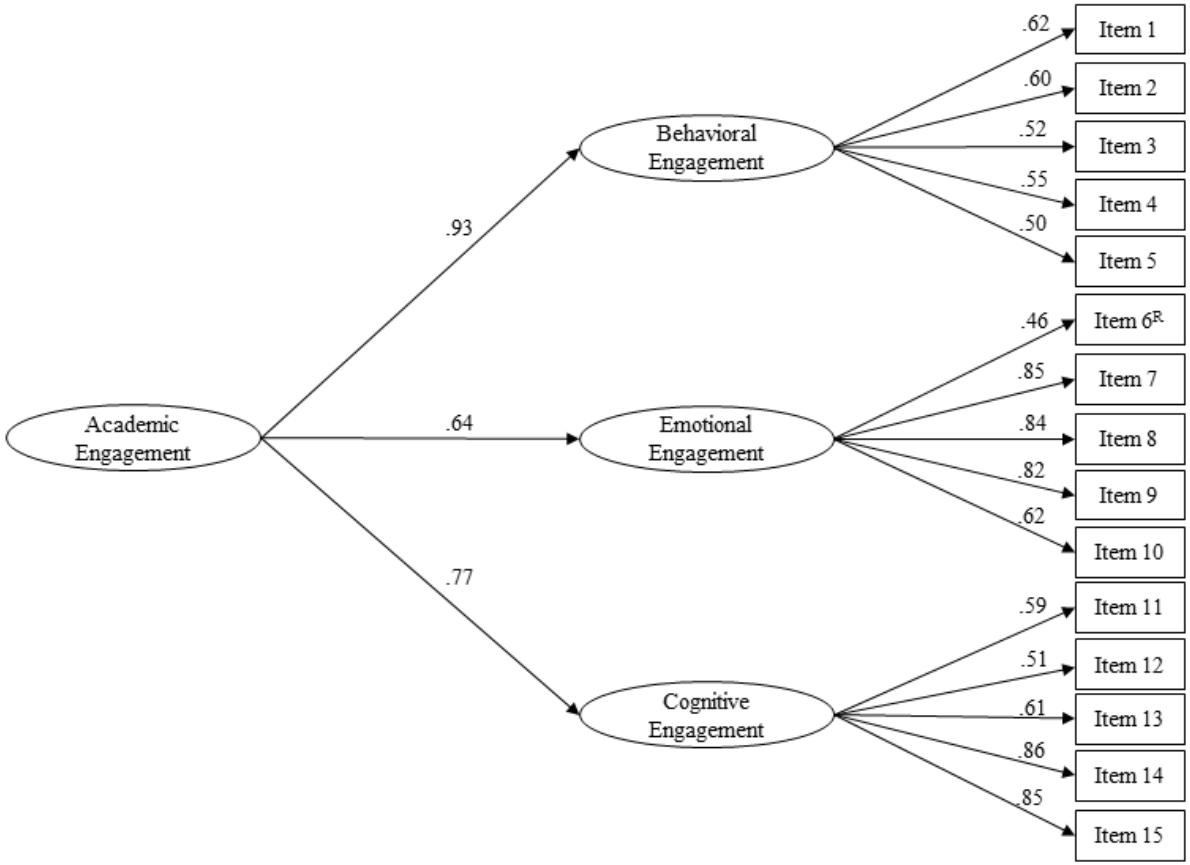
Notes. $^{ns}p > .05$; $^{*}\ p < .05$; $^{***}\ p < .001$

Table 4. USEI's models comparisons for scientific area of graduation

Model invariance	χ2	df	χ2/df	*CFI* scaled	Δχ2	Δ*CFI* scaled
Configural (factor structure)	539.91	262	2.06	.959	-	-
First-order loadings invariance	497.07	285	1.74	.969	13.14^{ns}	.010
Second-order loadings invariance	579.04	290	2.00	.962	25.70^{***}	.007
Thresholds of measured variables	684.93	373	1.84	.963	28.70^{ns}	.001
Intercepts of first-order factors invariance	736.91	379	1.94	.962	5.67^{ns}	.001
Disturbances of first-order factors invariance	747.34	384	1.95	.963	2.17^{ns}	.001
Residual variances of observed variables invariance	841.20	414	2.03	.961	28.32^{**}	.002

Notes. $^{ns}p > .05$; $^{**}p < .01$; $^{***}p < .001$


Behavioral Engagement
Emotional Engagement
Cognitive Engagement
.72
.60
.49
.62
.61
.52
.55
.50
.46
.85
.84
.82
.62
.59
.51
.61
.86
.85
Item 1
Item 2
Item 3
Item 4
Item 5
Item 6R
Item 7
Item 8
Item 9
Item 10
Item 11
Item 12
Item 13
Item 14
Item 15


Academic Engagement
.93
.64
.77
Behavioral Engagement
Emotional Engagement
Cognitive Engagement
.62
.60
.52
.55
.50
.46
.85
.84
.82
.62
.59
.51
.61
.86
.85
Item 1
Item 2
Item 3
Item 4
Item 5
Item 6R
Item 7
Item 8
Item 9
Item 10
Item 11
Item 12
Item 13
Item 14
Item 15