Analyzing measurement invariance of The Students' Engagement Instrument Brief version – The cases of Denmark, Finland, and Portugal

Virtanen, T. E.[1a] Moreira, P.,[2] Ulvseth, H.,[3] Andersson, H.,[4] Tetler, S.,[5] and Kuorelahti, M.[6]

[1] University of Jyvaskyla, Department of Teacher Education, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland; tuomo.e.virtanen@jyu.fi
[2] Lusíada University, R. Dr. Lopo de Carvalho, 4369-006 Porto, Portugal; paulomoreira@por.ulusiada.pt
[3] Aarhus University, The Danish School of Education, Tuborgvej 164, 2400 Copenhagen NV, Denmark; hildu@edu.au.dk
[4] Malmö University, Faculty of Education and Society, box 20506 Malmö, Sweden; helena.andersson@mah.se
[5] Aarhus University, The Danish School of Education, Tuborgvej 164, 2400 Copenhagen NV, Denmark; tetler@edu.au.dk
[6] University of Jyvaskyla, Department of Education, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland; matti.kuorelahti@jyu.fi

[1a] Corresponding author. Address: University of Jyvaskyla, Department of Teacher Education, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland. Tel.+358-40-8053636.
E-mail address: tuomo.e.virtanen@jyu.fi

The study has been supported by a grant from the Academy of Finland (No. 268586 fro 2013–2017)

Abstract

The promotion of students' engagement with school is an internationally acknowledged challenge in education. There is a need to examine the structure of the concept of student engagement and to discover the best practices for fostering it across societies. That is why the cross-cultural invariance testing of students' engagement measures is highly needed. This study aimed, first, to find the reduced set of theoretically valid items to represent students' affective and cognitive engagement forming the Brief-SEI (brief version of the Student Engagement Instrument; SEI; Appleton, Christenson, Kim, & Reschly, 2006). The second aim was to test the measurement invariance of the Brief-SEI across three countries (Denmark, Finland, and Portugal). A total of 4,437 seventh-grade students completed the SEI questionnaires in the three countries. The analyses revealed that of the total 33 original instrument items, 15 items indicated acceptable psychometric properties of the Brief-SEI. With these 15 items, cross-national factorial validity and invariances across genders and students with different levels of academic performance (samples from Finland and Portugal) were demonstrated. This article discusses the utility of the Brief-SEI in cross-cultural research and its applicability in different national school contexts.

Keywords: supporting students' engagement with school, cross-cultural student engagement research, Student Engagement Instrument, measurement invariance

Analyzing measurement invariance of The Students' Engagement Instrument Brief version – The Cases of Denmark, Finland, and Portugal

Student engagement refers to the students' subjective experiences connected to school, resulting in the sense of connectedness and identification with school (Fredricks, Blumenfeld, & Paris, 2004). Typically, it has been described as a three-dimensional model consisting of affective, behavioral, and cognitive dimensions (Fredericks et al., 2004). In short, highly engaged students experience school belonging and enjoy learning (affective dimension), participate in school activities (behavioral dimension), and are interested in academic content and use effective learning strategies (cognitive dimension). As student disengagement may be evidenced first at the psychological level and only thereafter in behavior (Archambault, Janosz, Fallu, & Pagani, 2009), affective and cognitive engagement are primary targets of interventions that aim to prevent student disengagement from school. In the present study, we saw student engagement as a three-dimensional construct (affective, behavioral, and cognitive dimensions) but focused on the psychological subtypes of engagement, namely affective and cognitive engagement.

The level of student engagement with school is one of the strongest predictors of academic trajectories, from short- to long-term outcomes (Archambault, Janosz, Fallu, & Pagani, 2009). Despite the very substantial developments in research on engagement in the last decade, the measurement invariance of instruments capturing student engagement across societies is still in its embryonic phase (Samuelsen, 2012). To conduct meaningful cross-cultural research on student engagement and develop engaging school practices that are internationally applicable, a pressing need exists for research examining the invariance of student engagement instruments across societies and educationally relevant groups, including genders and students with a variety of academic performance levels. Additionally, there are virtually no theoretically sound student engagement instruments serving both research and practice purposes. Optimally, such instruments

should be both valid, reliable, easily administered, quickly completed by students, and easily scored and interpreted by practitioners. These qualities are particularly important and challenging when measuring highly inferable affective and cognitive subtypes of engagement. One widely used instrument tapping lower secondary students' affective and cognitive engagement is the 33-item Student Engagement Instrument (SEI; Appleton et al., 2006). The validity of the SEI has been demonstrated in several studies describing its psychometric properties in different populations. The SEI has shown to be an adequate instrument to capture engagement of students from different schooling levels (Betts, Appleton, Reschly, Christenson & Huebner, 2010; Carter Reschly, Lovelace, Appleton & Thompson, 2012). Also, it has been shown to present good concurrent (Appleton et al., 2006; Lovelace, Reschly, Appleton & Lutz, 2014) and predictive validity (Lovelace et al., 2014) along with good reliability scores (Betts et al., 2010; Lovelace et al., 2014). Finally, there is preliminary evidence for the adequacy of the SEI in capturing dimensions of students' engagement with school in different societies (Moreira, Vaz, Dias, & Petracchi, 2009; Virtanen, Kiuru, Lerkkanen, Poikkeus, & Kuorelahti, 2016). Of special relevance would be the development of short version of the SEI. It would help professionals monitor students' engagement on a regular basis enabling them to detect early signs of student disengagement, and modify teaching practices accordingly. Having a Brief-SEI favors these processes because it is less time consuming, and it will reduce student fatigue and boredom associated with answering redundant items. A brief version does not necessarily mean less validity of the measure, as research shows that shorter versions of assessment instruments may be as valid as the longer versions. The objective of this study was to examine the measurement invariance of the 15-item Brief version of the Student Engagement Instrument across countries, genders, and students with different levels of academic performances. We chose the same grade, namely Grade 7, in the three countries representing different regions of Europe (two from Scandinavia and one from Southern Europe). We wanted to study the measurement invariance across countries and simultaneously have participants with

equivalent developmental features so that eventual differences on the Brief-SEI factorial structure would not be caused by students' developmental specificities. Moreover, research showed that students' engagement decrease as their school level increase (Christenson, Reschly, & Wylie, 2012). In order to have time to work with the students at risk, students in the beginning of secondary education were selected as study participants.

Danish educational system

Education is compulsory between the ages of 6 and 16 in Denmark. The Danish Folkeskole is a comprehensive school covering both primary and lower secondary education, that is, the first- (Grades 1 to 6) and second- (Grades 7 to 9/10) stage basic education. The vast majority of Danish students attend public schools, but approximately 17 percent attend a private school (Cevea, 2014). Students do not get grades or any other formal marks before Grade 8. Denmark is performing above average level in the international PISA (Programme for International Student Assessment) tests, which regularly measure 15-year-old students' scholastic performance in mathematics, science, and reading in the OECD (Organization for Economic Co-operation and Development) and some non-OECD countries (OECD, 2016). Furthermore, Danish students express general satisfaction with school (OECD, 2016).

Finnish educational system

The primary (Grades 1–6, from age 7 to 12) and lower secondary (Grades 7–9, from age 13 to 16) school form an integrated compulsory basic education, which lasts nine years. Virtually all schools are public, with small between-schools variation in student performance. There are no standardized tests, inspections, or school rating systems. Finland repeatedly performs at the top among European countries in the international PISA program (OECD, 2013; OECD, 2016). However, Finnish results have not been optimal regarding student engagement. For example, 13-

and 15-year-old students have reported low levels of liking school and high levels of pressure from schoolwork (Currie et al., 2012).

Portuguese educational system

Compulsory education is organized in terms of elementary school (4 years, from age 6), middle school (5 years), and secondary school (3 years). The great majority of schools are public schools. In 2009, the amount of mandatory education changed from 9 to 12 school years. Portuguese students performed about the same or better than the OECD average in PISA 2016 assessment depending on the subject (OECD, 2016). They, however, reported lower levels of behavioral engagement than the OECD average. Every third Portuguese student has reported "having skipped classes or days of school," while the average was one-fourth of pupils in the other European countries. However, Portuguese students reported motivation for studying in line with the average (OECD, 2013).

Measurement invariance of Student Engagement Instrument

Measurement invariance means that the measured construct carries universal features across groups representing, for example, different genders and cultures or respondents measured at different points of time (see Meredith & Teresi, 2006). With the measurement invariance across cultures and students with varying characteristics confirmed, the ultimate differences would not be the result of systematic measurement bias; instead, they are real differences on the given psychological construct. Establishing measurement invariance is a necessary condition for fair and equal selection procedures, when tests are to be used in screening of individuals for intervention (Borsboom, 2006).

Studies conducted in the U.S. have found evidence for the invariance of the two-dimensional (affective and cognitive) Student Engagement Instrument for some indicators. The SEI presents similar factor structure (the same number of factors and the similar patterns of factor loadings)

among students in Grades 6 to 12, genders (Betts et al., 2010) and different economic status student groups (Stafford, 2011).

Method

Participants and Procedure

Sample Denmark. The Danish sample consisted of 1295 Grade 7 students (49.9 percent girls) from 22 schools. The response rate was 95.1 percent. The SEI was translated by the researchers in cooperation with bilingual teachers, and it was modified into a Danish school context. Before the data collection, the questionnaire was piloted in two seventh-grade classrooms located in different areas in Denmark. The schools studied were selected to ensure geographical spread of the schools as well as variation in the schools' performances on national tests. All of the selected schools were public, all agreed to participate on a voluntary basis, and all distributed a letter to the students' parents or guardians in which the nature of the study was explained, along with the procedure for withdrawing their child from participation. The surveys were administered to students by the researcher.

Sample Finland. The Finnish sample consisted of 793 Grade 7 students (51.2 percent girls) from eight schools. The response rate was 86.3 percent. The SEI was translated into Finnish by a certified translator, and it was piloted in an urban lower secondary school in order to gain user feedback. Taking account of student feedback, the SEI was slightly modified and was back-translated into English; the translation was then compared with the original translation. The participating schools were typical public schools – five were located in northern Finland and three in western Finland. Following the guidelines of the Finnish National Advisory Board on Research Ethics (2009), the schools distributed a letter to the children's parents or guardians in which the nature of the study was explained, along with the procedure for withdrawing their child from participation. The surveys were administered to students by teachers.

Sample Portugal. The Portuguese sample consisted of 2405 seventh-grade students (52.4 percent girls) from 10 public schools in northern Portugal. Response rate was 85.1 percent. After the schools agreed to participate in the study, consent was collected from the students' parents or guardians. The mean age of seventh graders was 12.5, meaning that they were approximately one year younger than in the Scandinavian samples because the age of school start is one year earlier, at the age of six.

Measures

The Student Engagement Instrument (SEI; Appleton et al., 2006) is a 35-item, student-reported measure capturing students' engagement with school throughout two major dimensions: affective and cognitive engagement both consisting of three factors. Further studies (Betts et al., 2010; Reschly, Betts, & Appleton, 2014) have excluded a cognitive engagement factor–Extrinsic Motivation–because its reliability may be compromised due to the two reverse-scored items. The present study analyzed the 33-item SEI structure consisting of five interrelated factors. Three factors pertained to affective engagement: Teacher–Student Relationships (TS), Peer Support at School (PS), and Family Support for Learning (FS). Two pertained to Cognitive Engagement: Control and Relevance of School Work (CR) and Future Aspirations and Goals (FG). Affective engagement refers to students' experiences of contextual features of school in terms of connection with others and cognitive engagement to student-perceived relevance of school. The items were rated on a 4-point scale (1 = *strongly agree*; 4 = *strongly disagree*). The items were reverse-coded so that higher scores indicated a higher level of engagement.

Students' academic performance was available for only Finnish and Portuguese students because in Denmark, students do not get grades or any other formal marks before eighth grade. Due to the differences in Finnish and Portuguese grading systems, students' grades were standardized ($M = 0$, $SD = 1$), after which approximately three equal-size groups based on standardized values were formed: Good performers, Moderate performers, and Poor performers.

Students' gender was coded as 0 indicating boys and 1 indicating girls.

Data analysis strategy

Item selection was guided by statistical and conceptual considerations in an iterative process that continued until three items per factor were identified. The primary criterion for selecting the items was high-factor loadings on the corresponding factors. Secondarily, item pairs with high residual covariances were avoided. Third, items with redundant item contents were evaded (Hayduk & Littvay, 2012). The method of removing items capturing similar information has been used to abbreviate measures without significantly altering the meaning of the scale (Meads & Bentall, 2008). For example, choosing the items for Peer Support at School scale was based on the following logic. Items s6, s7, and s14 showed the highest loadings on Peer Support at School in the Portuguese sample. Of these items, s6 and s7 loaded highly on Peer Support at School in the Finnish sample and s14 in the Danish sample. Due to similar item contents, residuals of items s6 and s7 correlated highly with each other. Therefore, item 6 was replaced with item s4.

Testing for invariance entailed a hierarchical set of steps that began with the determination of a well-fitting multigroup configural model (combination of the three baseline models) for which sets of parameters were put to the test of invariance in a logically ordered and increasingly restrictive fashion (Byrne, 2008). The successive models (configural invariance model against weak factorial invariance model, weak factorial invariance model against strong factorial invariance model, and strong factorial invariance model against strict factorial invariance model) were compared against each other using Satorra-Bentler scaled chi-square (Bollen, 1989). If the more restrictive model did not fit significantly worse than the previous less restrictive model, the assumption of the higher-level invariance was supported. However, rejecting the null hypothesis of less restrictive model invariance precludes a test of higher-level invariance. Invariances across countries, genders, and different levels of students' academic achievement (Finnish and Portuguese samples) were tested.

The parameters were estimated using maximum likelihood with robust standard errors estimator (MLR; Muthén & Muthén, 1998–2012). The percentage of missing values varied between 0.5 percent (*Overall, adults at my school treat students fairly* in Danish sample) and 6.1 percent (*Students here respect what I have to say* in Finnish sample). As indicated by Little's MCAR test (Little, 1988), the missing values were not missing completely at random in all the samples: $\chi^2_{DEN}(2967) = 3115.47$; $p = .028$, $\chi^2_{FIN}(4926) = 4978.15$; $p = .298$, $\chi^2_{POR}(3960) = 4853.34$; $p < .001$. Therefore, the missing values were imputed with the M*plus* Bayesian multiple-imputation method (Rubin, 1987). This method averages the parameter estimates over the set of analyses (50 imputed data sets) and computes standard errors using the average of the standard errors over the set of analyses and the between-analysis parameter estimate variation. The analyses were conducted with M*plus* version 7.11 (Muthén & Muthén, 1998–2012). The goodness-of-fit of the estimated models was evaluated according to Root Mean Square Error of Approximation (RMSEA) and Comparative Fit Index (CFI). The cutoff values for well-fitting models were RMSEA < .05 and CFI > .95 (Byrne, 2012).

Results

The iterative process of item selection yielded the following items: items s16, s21, s31 for Teacher–Student Relationships; items s4, s7, and s14 for Peer Support at School; items s12, s20, and s29 for Family Support for Learning; items s33, s34, and s35 for Control and Relevance of School Work; and, finally, items s8, s19, and s30 for the Future Aspirations and Goals. The Brief-SEI items are listed in Appendix A.

Analysis of country-specific correlations between student engagement variables and academic achievement indicated that student engagement variables correlated positively and statistically significantly with each other. The magnitude of correlations was moderate except for the correlation between Future Aspirations and Goals and Control and Relevance, which was high (Cohen, 1988) in each country. Student engagement dimensions were interrelated while clearly describing distinct

aspects of student engagement. Moreover, student engagement dimensions correlated positively and statistically significantly with students' academic achievement (Finnish and Portuguese data sets) indicating some evidence for concurrent validity. The correlations between students' academic achievement and the dimensions of student engagement for Finnish and Portuguese data sets were as follows: Teacher-Student Relationships .19 and .07; Peer Support at School .07 and .09; Family Support for Learning .19 and .17; Control and Relevance of the School Work .27 and .16; and Future Aspirations and Goals .34 and .26. The correlations were consistent across the two countries and statistically significant at $p < .001$ except for the correlation between academic achievement and Teacher-Student Relationships in the Portuguese sample ($p < .01$). Higher levels of one student engagement dimension related to higher levels of student engagement in all the other dimensions and students' academic achievement. Portuguese students (M_{range} 3.15–3.70) reported the highest engagement followed by Danish (M_{range} 3.08–3.57) and Finnish students (M_{range} 2.86–3.32). Finnish students' ($M = 1.01$, $SD = 1.00$) academic achievement was better than their Portuguese counterparts' ($M = -0.33$, $SD = 0.75$): $t(1064.08) = 34.16$, $p < .001$.

Table 1 reports standardized factor loadings, scale reliabilities (Cronbach alphas), and correlations with the full scale for each country separately. Brief-SEI factor loadings were good for all 15 items in all three countries. Thirty-one out of 45 standardized first-order factor loadings exceeded .60 and 11 loadings exceeded .70. Only one item (s35 for Danish sample) had a factor loading less than .50 (.46). The second-order factor loadings were also high, exceeding .93. Scale reliability information indicates acceptable or good reliabilities with alphas varying between .55 and .83. The lack of excellent alpha coefficients may partly be caused by the reduced number of items (three per factor) and the Alpha's tendency to underestimate reliabilities. Moreover, the item selection aimed to capture slightly different aspects of the given conceptual variable. Therefore, similar item contents and wordings were avoided on purpose which most likely lowered the alpha coefficients. The factor correlations between the Brief-SEI and the SEI five-factor full scales (Betts

Table 1

Standardized Factor Loadings, Scale Reliabilities, and Correlations with the Full Scale

	Standardized factor loadings			Scale reliabilities Cronbach's Alpha			Brief scale correlation with the full scale		
	Den	Fin	Por	Den	Fin	Por	Den	Fin	Por
Teacher–Student Relationships				.71	.83	.68	.90	.91	.88
s16	.61	.77	.66						
s21	.67	.79	.61						
s31	.72	.80	.66						
Peer Support at School				.69	.77	.64	.92	.92	.91
s4	.68	.70	.58						
s7	.56	.73	.59						
s14	.74	.75	.67						
Family Support for Learning				.70	.71	.63	.97	.96	.96
s12	.61	.61	.58						
s20	.71	.72	.66						
s29	.69	.69	.58						
Control and Relevance of School Work				.60	.69	.61	.84	.88	.82
s33	.64	.69	.52						
s34	.62	.73	.68						
s35	.46	.58	.57						
Future Aspirations and Goals				.63	.71	.55	.94	.92	.86
s8	.60	.66	.55						
s19	.66	.71	.58						
s30	.53	.64	.50						

Table 2 continues

Table 2 continued

	Standardized factor loadings			Scale reliabilities Cronbach's Alpha			Brief scale correlation with the full scale		
	Den	Fin	Por	Den	Fin	Por	Den	Fin	Por
Cognitive engagement				.88	.91	.88	.92	.93	.89
Control and Relevance of School Work	.96	.93	.93	-	-	-	-	-	-
Future Aspirations and Goals	.98	.98	1.01	-	-	-	-	-	-

Note. Cognitive engagement is a second-order factor consisting of Control and Relevance of School Work and Future Aspirations and Goals. Den = Denmark, Fin = Finland, Por = Portugal.

et al., 2010) were high (correlation coefficients between .82–.97) indicating that briefing the scales did not lead to major loss in information.

Establishing baseline models. The first step in establishing the Brief-SEI measurement invariance was establishing the well-fitting baseline model for each country separately (Byrne, 2012). This stage revealed unwarranted measurement redundancy between Future Aspirations and Goals and Control and Relevance of School Work factors. The factor correlation was > .90 in each country. As a result, these two first-order factors were specified as second-order factors (Cognitive Engagement) indicators. As seen in Table 2, baseline models fitted well in the data sets in each country (all RMSEA's $\leq$ .04 and CFI's $\geq$ .95). There were no correlation residuals greater than > .10 (Kline, 2011) indicating that model correlations were well explained by the data sets. The baseline models' Chi square values were significant ($p < .001$) which is likely due to the large sample sizes (see Brown, 2006). Moreover, baseline models for boys and girls and three students' academic performance groups (not reported in Table 2) were a good fit (for both genders RMSEA = .03 and CFI = .97; for academic performance groups RMSEA's $\leq$ .04 and CFI's $\geq$ .96). The Brief-SEI is presented in Figure 1.

Table 2

Fit Indices of the Estimated Models

Model	χ^2	*df*	*p*	RMSEA	CFI	$\Delta\chi^2(df)$	*p*
Baseline models							
Denmark	187.58	82	< .001	.03	.97		
Finland	170.96	82	< .001	.04	.97		
Portugal	355.78	82	< .001	.04	.95		
Configural invariance: Factor loadings and item intercepts estimated	717.91	246	< .001	.04	.96		
Weak factorial invariance: Factor loadings equal, item intercepts estimated	747.55	268	< .001	.04	.96	31.69(22)	.083
Strong factorial invariance: Factor loadings equal, item intercepts equal	2,471.40	300	< .001	.07	.81	1872.80(32)	<. 001

Note. RMSEA = Root Mean Square Error of Approximation. CFI = Comparative Fit Index.

Testing configural invariance. The next step was to specify a configural model for each of the following: the three countries, genders, and the three groups with different levels of academic achievement. Configural models incorporated all the baseline models and served as baseline models for comparing subsequent restricted models. In order to ensure the model identification, factor means were fixed at zero. Factor loadings and item intercepts were freely estimated for each group. All the configural models showed good fit to the data ($RMSEA_{countries}$ = .04 and $CFI_{countries}$ = .96, $RMSEA_{genders}$ = .03 and $CFI_{genders}$ = .97, $RMSEA_{achievement\ groups}$ = .03 and $CFI_{achievement\ groups}$ = .97). This result indicated that the number of factors and the pattern of item loadings were similar across the three countries, genders, and students with different levels of academic performances. Consequently, configural invariance was achieved.

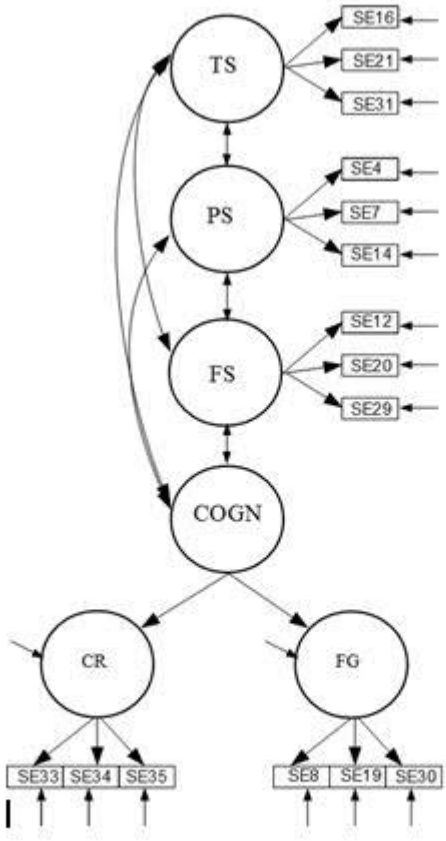

Figure 1. The Brief-SEI.

Testing for weak factorial invariances. Next, we specified three factor invariance models with the first- and second-order factor loadings freely estimated for the first data set and constrained to equity for the other data sets. Factor means were fixed at zero while the item intercepts were estimated. Factor loadings invariance model fitted the data well for the countries ($RMSEA_{countries}$ = .04 and $CFI_{countries}$ = .96), for genders ($RMSEA_{genders}$ = .03 and $CFI_{genders}$ = .97), and for student groups with different academic levels ($RMSEA_{achievement\ groups}$ = .03 and $CFI_{achievement\ groups}$ = .97). Difference testing revealed that the restrictions on the factor loadings to equity did not worsen the model across countries ($\Delta\chi^2$ (22) = 31.69, p = .083), genders ($\Delta\chi^2$ (11) = 16.56, p = .122), and students with different levels of academic performances ($\Delta\chi^2$ (19) = 19.33, p = .436). The invariance

of the factor loadings for the three countries, genders, and students with different academic levels were, therefore, confirmed.

Testing for strong factorial invariance. The equality restrictions were next extended to the item intercepts. First-order factor means were freely estimated in all but the reference group, while second-order factor mean was fixed at zero in all groups. The invariance of the factor loadings and item intercepts yielded models that were statistically significantly worse than the invariant factor loadings models (each Satorra-Bentler chi square difference test yielded $p <. 001$) (see Table 2). The strong factorial invariance of the Brief-SEI was, therefore, not demonstrated across countries, genders, and students with different levels of academic performances. As a consequence, the strict factorial invariance (invariant residual variances) was not tested.

Discussion

There is an increasing need to develop theoretically and psychometrically sound and easy-to-use measurement instruments for practitioners to assess student engagement with school. At the same time, there is a call for international studies aimed at providing a broader understanding of student engagement in various cultures. The present study had two aims. First, we aimed to find the reduced set of theoretically sound items to represent students' affective and cognitive engagement forming the brief version of the Student Engagement Instrument (SEI; Appleton et al., 2006). Second, we aimed to expand the evidence of the measurement invariance of the Brief-SEI across seventh graders in three countries, namely Denmark, Finland, and Portugal.

First, the results showed that the factorial structure of the brief version of the Student Engagement Instrument differed from that of the full scale. Instead of the five (Betts et al., 2010) or six (Appleton et al., 2006) interrelated first-order factors, the Brief-SEI was construed along three first-order factors forming affective engagement and one second-order factor (Cognitive Engagement) consisting of two first-order factors. Control and Relevance of the School Work and Future Aspirations and Goals factors were statistically inseparable; therefore, they were specified as

indicators measuring Cognitive Engagement. This is in line with previous studies that have found the Student Engagement Instrument cognitive engagement to be problematic in some samples (see AUTHOR, 2016; Carter, Reschly, Lovelace, Appleton, & Thompson, 2012; Grier-Reed, Appleton, Rodriguez, Ganuza, & Reschly, 2012).

Second, the current study demonstrated the configural and factorial invariances of the 15-item Brief-SEI. The number of factors, loading patterns, and factor loadings were equivalent in each country. The Brief-SEI fit very well in the data sets collected in the three countries, evidencing good construct validity. The reliability information indicated acceptable reliabilities. The correlations between the reduced scales and the full scales were in general high, implying that there was no major loss of information when using the Brief-SEI as compared to the 33-item SEI. Moreover, the Brief-SEI correlations between students' academic achievement and the dimensions of student engagement were, although generally small in magnitude, in line with previous studies (e.g., Appleton et al., 2006) showing that better academic achievement associated with higher engagement. In all, the current analyses indicated that the Brief-SEI has a potential to invariantly describe complex and dynamic interactions between individuals' characteristics and contextual influences across countries, genders, and students with different levels of academic performances.

However, the analysis of equal factor loadings simultaneously with equal item intercepts across the three countries did not verify the strong measurement invariance. In such cases item mean comparisons across groups may not be valid (Meredith & Teresi, 2006). Not achieving strong measurement invariance in the present study may stem from two kinds of method biases (see van de Vijver & Tanzer, 2004). First, there were minor differences in data collection. The Danish data were collected by the researcher; therefore, the survey administration was similar in all classrooms. The Finnish data were collected by the teachers who followed the instructions given by the researcher, so that there might have been variance in the way students were instructed to fill in the questionnaires. Second, there may have been cultural differences in students' response styles

(Wetzel, Carstensen, & Böhnke, 2013). This conjecture is supported by the notion that despite Finnish students' high PISA rankings (OECD, 2013), the lowest scores in the engagement scales were systematically found among them, implying a different cultural approach for responding to any questionnaire (i.e., systematically avoiding extreme ratings on the high end of the scales). Third, outside the method biases, Portuguese seventh graders were a year younger than their Scandinavian counterparts due to the differences in their school systems. This may well be a source of variance. However, the study showed that the concept of student engagement with school carries cross-cultural invariant elements, thus supporting earlier findings in North America (Wang et al., 2011).

Practical implications

The Brief-SEI can be used for student engagement research purposes. Established configural and factorial invariances allow cross-cultural comparisons in terms of latent construct structural parameters, including variances, covariances, and path coefficients (Cheung, 2008). For example, it is meaningful to compare the relationships between the constructs and the magnitude of path coefficients of the three dimensions of affective engagement on the outcome of interest. Furthermore, the 15-item brief version of the Student Engagement Instrument can offer a deeper understanding of highly interpretative subtypes of student engagement to the school personnel, including school psychologists and teachers, thus forming a basis for discussions with students and parents that aim to create more engaging learning environments for all students. Brief-SEI is easily administered by practitioners and quickly completed by students. It can be systematically and routinely used for data-based, school-level monitoring of student affective and cognitive engagement. Based on the Brief-SEI measures, low-engaged students and classrooms could be identified, and targeted interventions should be focused on them. Particularly, an affective engagement scale offers clear intervention targets for individual students and classrooms showing collective low engagement. This enables the establishment of the assessment-to-intervention link,

where students showing warning signs of low engagement can be identified at early phases and offered timely interventions (see Appleton, 2012).

Limitations and future directions

There are limitations to the study. First, the study data sets did not allow the invariance testing for other than seventh graders. The seventh grade is the first grade after transition from primary to lower secondary education; therefore, it is the time for gradual decline in various indicators of adolescents' academic motivation (Eccles & Midgley, 1989) and engagement (Skinner, Furrer, Marchand, & Kindermann, 2008). The Brief-SEI invariance testing should be extended to other lower secondary grades as well. Second, the Brief-SEI construct validity was evidenced in the three countries, but information concerning the concurrent validity relies only on one indicator, namely the relationship between engagement dimensions and academic achievement in Finnish and Portuguese samples. Further studies should examine the associations between the Brief-SEI factors and other educationally relevant constructs such as students' behavioral engagement, motivation, and school-related burnout. Finally, this cross-national study was conducted in three European countries. The applicability and generalizability of the Brief-SEI should also be studied in North America where the 33-item SEI is already largely used by practitioners and researchers. The already existing data sets could be used to test the functionality of the Brief-SEI in North America. Our study is a step in the dynamic process of cross-cultural research on students' engagement with school. It aims to inspire further studies on the applicability of the Brief-SEI in the North American school context. This would prompt advances on students' engagement with school both from the perspectives of practice and research.

References

Appleton, J. J. (2012). System consultation: Developing the assessment-to-intervention link with the student engagement instrument. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 725–741). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4614-2018-7_35

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. http://dx.doi.org/10.1016/j.jsp.2006.04.002

Archambault, I., Janosz, M., Fallu, J., & Pagani, L. S. (2009). Student engagement and its relationship with early high school dropout. *Journal of Adolescence, 32*(3), 651–670. http://dx.doi.org/10.1016/j.adolescence.2008.06.007

AUTHOR, 2016.

Betts, J. E., Appleton, J. J., Reschly, A. L., Christenson, S. L., & Huebner, E. S. (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. http://dx.doi.org/10.1037/a0020259

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

Borsboom, D. (2006). When does measurement invariance matter? *Medical Care*, *44*(11 Suppl 3), S176-S181.

Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, *20*(4), 872–882.

Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York, NY: Routledge. http://dx.doi.org/10.4324/9780203807644

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. http://dx.doi.org/10.1037/a0029229

Cevea (2014). Mere end hvert sjette barn går på privatskole [more than every sixth child attends private school]. Retrieved from http://www.cevea.dk/analyse/end-hvert-sjette-barn-gaar-paa-privatskole

Cheung, G. W. (2008). Testing equivalence in the structure, means, and variances of higher-order constructs with structural equation modeling. *Organizational Research Methods, 11*(3), 593–613. http://dx.doi.org/10.1177/1094428106298973

Christenson, A. L. Reschly, & C. Wylie (2012). *Handbook of research on student engagement.* New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4614-2018-7_39

Currie, C., Zanotti, C., Morgan, A., Currie, D., de Looze, M., Roberts, C., et al. (Eds.). (2012). *Social determinants of health and well-being among young people. Health behaviour in school-aged children (HBSC) study: International report from the 2009/2010 survey. Copenhagen, WHO regional office for Europe* (health policy for children and adolescents, no. 6).

Eccles, J. S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. In C. Ames & R. Ames (Eds.), *Research on motivation in education.3, goals and cognitions* (pp. 139–186). San Diego, CA: Academic Press.

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. http://dx.doi.org/10.3102/00346543074001059

Grier-Reed, T., Appleton, J., Rodriguez, M., Ganuza, Z., & Reschly, A. L. (2012). Exploring the student engagement instrument and career perceptions with college students. *Journal of Educational and Developmental Psychology, 2*(2), 85–96. http://dx.doi.org/10.5539/jedp.v2n2p85

Hayduk, L. A., & Littvay, L. (2012). Should researchers use single indicators, best indicators, or multiple indicators in structural equation models? *BMC Medical Research Methodology, 12*(1), 159–175. http://dx.doi.org/10.1186/1471-2288-12-159

Little, R. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198–1202. http://dx.doi.org/10.2307/2290157

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. http://dx.doi.org/10.1177/0734282914527548

Meads, D. M., & Bentall, R. P. (2008). Rasch analysis and item reduction of the hypomanic personality scale. *Personality & Individual Differences, 44*(8), 1772–1783. http://dx.doi.org/10.1016/j.paid.2008.02.009

Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care, 44*(11 Suppl 3), S69-S77. http://dx.doi.org/10.1097/01.mlr.0000245438.73837.89

Moreira, P.A.S., Machado Vaz, F.; Dias, P., & Petracchi, P. (2009). Psychometric properties of the

Portuguese version of the School Engagement Instrument. *Canadian Journal of School Psychology, 24*(4), 303-317. http://dx.doi.org/10.1177/0829573509346680

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

National Advisory Board on Research Ethics. (2009). Ethical principles of research in the humanities and social and behavioural sciences and proposals for ethical review. Retrieved from http://www.tenk.fi/sites/tenk.fi/files/ethicalprinciples.pdf

OECD. (2013). Programme for international student assessment (PISA). Retrieved from http://www.oecd.org/pisa/

OECD. (2016). Programme for international student assessment (PISA). Retrieved from http://www.oecd.org/pisa/

Reschly, A. L., Betts, J. E., & Appleton, J. J. (2014). An examination of the validity of two measures of student engagement. *International Journal of School & Educational Psychology, 2*(2), 106-114.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* New York, NY: John Wiley & Sons. http://dx.doi.org/10.1002/9780470316696

Samuelsen, K. M. (2012). Part V commentary: Possible new directions in the measurement of student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 805–811). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4614-2018-7_39

Stafford, R. (2011). Evaluation of the Student Engagement Instrument: Measurement invariance across economic status and association with academic achievement (master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1502352)

van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue Europeenne De Psychologie Appliquee, 54*(2), 119–135. doi:10.1016/j.erap.2003.12.004

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), 136–157.

Wetzel, E., Carstensen, C. H., & Böhnke, J. R. (2013). Consistency of extreme response style and nonextreme response style across traits. *Journal of Research in Personality, 47*(2), 178–189. doi 10.1016/j.jrp.2012.10.010

Appendix A

Brief-SEI items

	Teacher–Student Relationships
s16	Overall, my teachers are open and honest with me.
s21	Overall, adults at my school treat students fairly.
s31	At my school, teachers care about students.
	Peer Support at School
s4	Other students here like me the way I am.
s7	Students at my school are there for me when I need them.
s14	Students here respect what I have to say.
	Family Support for Learning
s12	When something good happens at school, my family/guardian(s) want to know about it.
s20	When I have problems at school, my family/guardian(s) are willing to help me.
s29	My family/guardian(s) want me to keep trying when things are tough at school.

	Control and Relevance of School Work
s33	Learning is fun because I get better at something.
s34	What I'm learning in my classes will be important in my future.
s35	The grades in my classes do a good job of measuring what I'm able to do.
	Future Aspirations and Goals
s8	My education will create many future opportunities for me.
s19	School is important for achieving my future goals.
s30	I am hopeful about my future.