Running Head. Student Engagement, Externalizing and Internalizing Behaviors

Internalizing and Externalizing Behavior Problems and Student Engagement in Elementary and Secondary School Students

Elizabeth Olivier[1†*], Alexandre J.S. Morin[1†], Jessika Langlois[2], Kristel Tardif-Grenier[3], & Isabelle Archambault[2]

[1]Substantive-Methodological Synergy Research Laboratory, Department of Psychology, Concordia University, Canada
[2]École de psychoéducation, Université de Montréal, Canada
[3]Département de psychoéducation et de psychologie, Université du Québec en Outaouais, Canada

[†]Because the first two authors (E.O. & A.J.S.M.) contributed equally to the preparation of this paper, their order was determined at random: both should be considered first author.

*Corresponding author
Elizabeth Olivier
Department of Psychology, Concordia University, Canada
7141 Sherbrooke W., Montreal, QC, Canada, H4B 1R6
(+1)514-848-2424 ext. 3633

Funding: The first author was supported by funding from the Canadian Institute for Military and Veteran Health Research (CIMVHR). The second author was supported by a grant from the Social Sciences and Humanities Research Council of Canada (435-2018-0368). Data collection was supported by grants from the Fonds de recherche Société et culture (elementary school project: #131430; secondary school project: 2018-NP-205005) and from the Social Sciences and Humanities Research Council of Canada (secondary school project; 430-2018-00456). The fifth author was also supported by funding from the Canada Research Chairs program.

This is the prepublication version of the following manuscript:
Olivier, E., Morin, A.J.S., Langlois, J. Tardif-Grenier, K., et Archambault, I. (2020). Internalizing and externalizing behavior problems and student engagement in elementary and secondary school. *Journal of Youth and Adolescence*.
Early view: https://doi.org/10.1007/s10964-020-01295-x

Abstract

Students with externalizing or internalizing behavior problems are at increased risk of underachievement and school non-completion, often due to their lower school engagement. Two studies were undertaken to assess the unique and joint (i.e., interactive) associations between behavior problems and engagement during two developmental periods; childhood and adolescence. These studies also aim to disentangle the contribution of global (externalizing and internalizing) and specific (hyperactivity/inattention, opposition/defiance, anxiety, depression) behavior problems on the global and specific aspects of student behavioral, emotional, and cognitive engagement. Study 1 was conducted among a sample of elementary school students (n=1,036; 3rd to 6th grade; mean age = 9.94 y.o.; 47.30% female; majority native Canadians) and Study 2 was conducted in secondary school (n=1,011; 7th and 8th grade; mean age = 12.93 y.o.; 55.77% female; 60.64% from immigrant background). Results of the bifactor-CFA and path analyses from both studies indicate that global externalizing behaviors were associated with lower global and specific behavioral engagement. In Study 1, global internalizing behaviors were also associated with lower global and specific cognitive engagement, whereas specific anxiety was associated with lower global and specific emotional engagement. In Study 2, specific depressive symptoms were associated with lower global and specific emotional engagement. Together, these two studies suggest that externalizing behaviors remain risk factors for student disengagement during childhood and adolescence, but that the risk posed by internalizing behaviors increases in importance for older students.

Keywords: externalizing behaviors; internalizing behaviors; student engagement; bifactor-CFA

Equity in education is an international concern (OECD, 2018). In North America, disparities in upper secondary completion rates reach up to 30% between students without and with externalizing and internalizing behaviors (MEES, 2019). This source of disparity is larger than that between students according to their sex (up to 7%), SES background (up to 20%), and ethnicity (up to 20%) (UNESCO, 2020). Moreover, up to one-third of students enrolled in regular schools display some level of externalizing (i.e., hyperactivity, inattention, opposition, and defiance) or internalizing (i.e., anxiety and depressive symptoms) behavior adjustment problems (Olivier et al., 2018). Student disengagement in school is among the main mechanism leading to student academic failure and dropout throughout schooling (Fredricks et al., 2016), and even more strongly for youth with internalizing or externalizing behaviors (Gut et al., 2012). However, it remains unclear which of these behavior problems are more damaging for student behavioral, emotional, and cognitive engagement, as well as if the damages they contribute to are different for boys and girls, and at different developmental periods (Al-Hendawi, 2012). Thus, the current study proposes to assess the nexus between behavior problems and student engagement at two developmental periods—before and after the transition to secondary school—and to assess whether some of the relations vary as a function of students' sex.

Student Behavior Problems: The Roles of Sex and Development

Students with behavior adjustment problems are often categorized as displaying externalizing or internalizing behaviors (e.g., Caci et al., 2015). Externalizing behaviors include an array of noticeable and disruptive behaviors for one's surroundings (Achenbach & Edelbrock, 1978). Behaviors classified as externalizing typically encompass hyperactivity, attention problems, and conduct problems such as opposition and defiance (Achenbach & Edelbrock, 1978). In contrast, internalizing behaviors are not as noticeable, as they pertain to more interiorized self-directed difficulties (Achenbach & Edelbrock, 1978). Behaviors classified as internalizing encompass worry, sadness, anhedonia (lack of interest and pleasure), and other negative moods and emotions. Even if externalizing and internalizing behaviors are considered to be distinct (Caci et al., 2015), some children (Willner et al., 2016) and adolescents (Reitz et al., 2005) will still go on to develop both types of problems (Papachristou & Flouri, 2019).

Although externalizing and internalizing behaviors often take root early in development, these two types of behavior follow distinct developmental trends. Externalizing behaviors are usually first recognized as problematic when students enter the school system and have to confront themselves with school and classroom expectations (Campbell et al., 2014). Partly because of the associated impaired self-regulation skills, hyperactive and inattentive behaviors are incompatible with expectations to pay attention to teachers' explanations, work quietly, and interact in a prosocial manner with their peers (Campbell et al., 2014). Similar difficulties have also been observed in children with oppositional and defiant behaviors (Bierman & Sasser, 2014). Such externalizing behaviors tend to persist and remain stable over time, from childhood into adolescence, especially for children who first present externalizing symptoms early in development (Bierman & Sasser, 2014). Although some students develop defiant and oppositional behaviors later in adolescence, most adolescents who display them already presented some externalizing behaviors in childhood (Bierman & Sasser, 2014). Moreover, boys are more numerous than girls to display externalizing behaviors, and this sex-related disparity in prevalence remains similar across development (opposition/defiance: Bierman & Sasser, 2014; hyperactivity/inattention: Campbell et al., 2014).

Developmental trends are slightly different for internalizing behaviors, especially between boys and girls. Anxious and depressive thoughts and emotions are less prevalent in childhood than in adolescence. Yet, when they emerge early in development, depressive symptoms (Garber & Rao, 2014) and anxiety (Vasey et al., 2014) tend to persist over time. During childhood, boys and girls present a similar risk of internalizing behaviors (depressive symptoms: Garber & Rao, 2014; anxiety: Vasey et al., 2014). However, with the arrival of puberty and adolescence, rates of internalizing behaviors, especially depressive symptoms, increase among boys and girls (Nivard et al., 2017). This increase is particularly marked for girls, resulting in higher levels of anxiety and depressive symptoms among adolescent girls than boys (Garber & Rao, 2014). Such developmental trends and sex differences in terms of internalizing and externalizing behaviors require further investigations to understand how behavior problems may lead students to experience negative school-related outcomes, such as lower levels of engagement (Eccles & Roeser, 2009).

Finally, co-occurring externalizing and internalizing problems may be more problematic when problems start early in the development (Nivard et al., 2017). Such co-occurring problems also tend to be more stable over youth development (Willner et al., 2016). However, it is unclear whether co-occurring internalizing and externalizing behaviors are more salient in boys or girls, as studies on this subject have yielded mixed results. Indeed Olivier et al. (2018) found that boys were more numerous to display externalizing behaviors, whereas girls were more numerous to display internalizing behaviors. Papachristou and Flouri (2019) found that boys were more likely to present higher levels of externalizing and internalizing behaviors. Finally, Reitz et al.'s (2005) results show that girls have higher levels of internalizing behaviors but that boys and girls do not differ in terms of externalizing behaviors. Despite the general recognition that externalizing and internalizing behaviors are independent risk factors, it is generally accepted that an accumulation of such problems is likely to bring about

additional risks (Evans et al., 2013). As youth spend an important proportion of their time ins school, these risks are likely to threaten their adjustment to school, particularly their engagement.

Student Engagement: Components and Developmental Trends

Student engagement reflects students' underlying motivation for their schoolwork (Skinner et al., 2009). As such, engagement is considered to be "motivation in action" (Skinner et al., 2009). According to the widely accepted three-dimensional definition proposed by Fredricks et al. (2004; see also Fredricks et al., 2016), student engagement encompasses behavioral, emotional, and cognitive components, all of which are necessary for active student investment in their schoolwork. Behavioral engagement refers to students' conduct and actions, including active participation, attendance, as well as compliance with classroom rules and teacher instructions (Fredricks et al., 2004). Emotional engagement encompasses students' affective reactions to the learning process and classroom environment, including their interest, enjoyment, and happiness (Fredricks et al., 2004). Cognitive engagement includes students' thought processes that promote their dedication of efforts to learning and mastery of school-related tasks and content. As such, cognitive engagement encompasses self-regulated and deep-processing strategies, such as reformulating ideas in one's own words, identifying important information, and trying different strategies to solve problems (Fredricks et al., 2004). Student engagement seems to remain stable across development for a majority of students, both in children (Archambault & Dupéré, 2016) and adolescents (Wang & Eccles, 2012). In addition, girls tend to display slightly higher engagement levels than boys, regardless of the developmental period considered, particularly for language classes (Archambault, Janosz, Morizot, et al., 2009). Apart from these different levels of engagement, boys and girls seem to follow similar developmental trends (Li & Lerner, 2011).

Behavior Problems and Student Engagement: Stage-Environment Fit Theory

According to Stage-Environment Fit Theory (Wigfield & Eccles, 2000), students display an optimal engagement when the school environment (the school policies, pedagogical practices, climate, etc.) offers a level of support that matches their developmental needs for belongingness, competence, autonomy, and security at each developmental stage. For most students, the fit between their school settings and their own needs as developing individuals is satisfying enough for them to maintain adequate engagement levels, leading to academic success. Students suffering from externalizing and internalizing problems usually need intensified support from their schools to display a level of engagement comparable to that of their peers. Yet, the support required for these youth is not always within reach of what schools can offer. Besides, despite this greater need, these students might end up receiving less support than their peers. Some of their behaviors can interfere with efficient classroom-functioning and their own individual functioning (Campbell et al., 2014). As such, students with externalizing and internalizing problems are likely to experience a greater level of stage-environment mismatch at school, possibly leading to lower school engagement levels.

Given that school resources are limited identifying which types of adjustment problems are the most important risk factors for students' school engagement might help better targeting students with the most pressing needs at different developmental periods. Moreover, these risks are likely to change at different stages of students' schooling, and possibly between boys and girls. The next section reviews studies assessing the nexus between adjustment problems and student engagement by contrasting the contribution of externalizing and internalizing behaviors, during childhood versus adolescence, and between boys and girls.

Behavior Problems and Student Engagement: An Empirical Review

Table 1[1] reviews previous studies that have focused on the school engagement of students displaying varying levels of externalizing and internalizing behaviors. A general overview of this table first depicts that studies assessing externalizing behaviors mainly focus on childhood, whereas studies assessing internalizing behaviors mainly focus on adolescence. When assessed separately from internalizing behaviors, externalizing behaviors have been systematically found to be associated with lower levels of behavioral engagement in children (see Table 1: Archambault et al., 2017; Demaray & Jenkins, 2011; Junod et al., 2006; Moilanen et al., 2010; Olivier & Archambault, 2017; Plamondon & Martinussen, 2019; Volpe et al., 2006) and adolescents (Tian et al., 2017; Wang & Fredricks, 2014). Despite these consistent findings, the relations between externalizing behaviors

[1] The literature review was conducted using the PsycInfo and Web of Science search engines as well (keywords details: (anxi* or depress* or internali* or hyperac* or inatten* or aggress* or defian* or opposit* or disrupt* or externali*) and (engag* or motiva*) and (school or student or classroom or acad*) which were restricted to the title and peer reviewed articles). The search yield a total of 234 article. Of those, 63 were retained based on the title, and 20 matched the inclusion criterial, namely to assess at least one behavior problem in association student engagement or a closely related measure (motivation achievement goals (n = 4), pre-school learning skills (n = 3), and school connectedness (n = 1) are excluded from Table 1, but discussed in the text when appropriate). Studies that assessed a subject-specific adjustment problem (e.g., mathematic anxiety) or focusing solely on delinquency involving illegal actions (e.g., bringing drugs to school) or outside of school delinquency (e.g., gang participation) were also excluded from Table 1. Screening the reference list of the retained article (i.e., forward snowballing) yield an additional 3 studies also reported in Table 1.

and the emotional and cognitive components of student engagement are far less clear. For instance, Archambault et al. (2017) found that students with high levels of opposition/defiance reported low emotional engagement, but Olivier and Archambault (2017) found no such association for students displaying hyperactivity/inattention. Among secondary school students, studies have considered emotional and cognitive engagements in a global measure of student engagement and found that hyperactive-inattentive and opposition/defiance behaviors led to a lower global engagement (see Table 1: Nguyen et al., 2019; Tian et al., 2017; Wang & Fredricks, 2014). Finally, the few studies that have assessed possible sex differences found that externalizing behaviors impacted boys' and girls' engagement similarly (see Table 1: Archambault et al., 2017; Demaray & Jenkins, 2011; Olivier & Archambault, 2017).

Most research focusing on the association between internalizing behaviors and student engagement has been conducted among samples of adolescents and has focused primarily on the effects of depressive symptoms. Among children, Kurdi and Archambault (2020) found that anxiety does not directly lead to decreased behavioral, emotional, and cognitive engagement, neither in boys nor in girls. In adolescence, studies tend to report that depressive symptoms negatively impact students' behavioral, emotional, or global engagement (see Table 1: Dorio et al., 2019; Fiorilli et al., 2017; Garvik et al., 2014; Wang & Peck, 2013; Wang et al., 2015). Although Dorio et al. (2019) found that depressive symptoms lead to a larger decrease in global engagement for girls, Derdikman-Eiron et al. (2011) instead found that internalizing behaviors, including anxiety and depressive symptoms, led to a lower behavioral engagement for all youth, but more pronounced for boys. Thus, there is no consensus on a sex that might be more at risk. In general, Garvik et al. (2014) warn that the effect sizes of depressive symptoms remain small, suggesting that even if internalizing behaviors seem to hamper students' engagement, most of them appear to be able to manage their symptoms to maintain adequate engagement. However, as for externalizing behaviors, associations involving the emotional and cognitive component of engagement are not as well established as those involving the behavioral dimension across developmental levels and samples of boys and girls.

Six studies have assessed the simultaneous repercussions of externalizing and internalizing behaviors on student engagement. Among children, Searle et al. (2013) found that hyperactivity-inattention, but not opposition-defiance and depressive symptoms were associated with lower global engagement. In contrast, Baker et al. (2008) found that internalizing behaviors, but not externalizing behaviors, led to lower behavioral engagement. Curhan et al. (2020) found similar results for adolescents' emotional engagement. Li et al. (2008) and Li and Lerner (2011), more specifically, studied opposition/defiance and depressive symptoms among adolescents. They found that these two adjustment problems led to decreases in behavioral and emotional engagement. Finally, Olivier et al., (2018; 2020) found that an externalizing profile of problems led to lower behavioral, emotional, and cognitive engagement for girls but not for boys. Girls with an internalizing profile of problems also reported a lower engagement. Yet, boys, but not girls, were likely to display externalizing and internalizing problems simultaneously, which led them to report a lower behavioral and emotional engagement. This last result suggests that there may be negative repercussions of accumulating behavior problems. For instance, Eisenberg et al. (2009) and Wang, F.L. et al. (2016) assessed students' self-regulatory abilities, a component of cognitive engagement. They found that youth displaying externalizing behaviors, or a combination of externalizing and internalizing behaviors, were more at risk of showing lower levels of such abilities than youth reporting only internalizing behaviors, or no problem. Such results call for a more in-depth assessment of the combined, and possibly harmful, influence of externalizing and internalizing behaviors on student engagement.

Finally, although the studies reported in Table 1 have alternatively relied on subject-specific (i.e., math or language) measures of engagement, this methodological consideration does not seem to have impacted the results. This observation could be partly explained by all of these studies having relied on domain-general (rather than subject-specific) measures of behavioral problems. In addition, although student engagement levels may differ across subjects, subject-specific engagement seems to remain importantly driven by motivation, affects, and engagement encompassing several school subjects (e.g., Gogol et al., 2017).

Toward a Better Understanding of the Global and Specific Components of Externalizing Behaviors, Internalizing Behaviors, and School Engagement

A precise assessment of the unique and combined effects of distinct behavior problems on student's engagement involves accounting for the high rate of co-occurrence between ratings of hyperactivity/inattention and opposition/defiance (Campbell et al., 2014), as well as between ratings of anxiety and depressive symptoms (Vasey et al., 2014). This high level of co-occurrence makes it difficult to precisely account for the role of global levels of externalizing/internalizing behaviors, relative to the role of specific behavior problems. A typical approach would be to focus on global levels of externalizing/internalizing behaviors, thus neglecting the possible role of specific types of behavioral problems. An alternative approach could focus on specific behavioral problems, which then carries the risk of obtaining results tainted by the role of the unmodelled global levels of externalizing/internalizing behaviors. Indeed, multivariate analyses estimate the role played by each predictor net of what it shares with the others. When these analyses involve highly related predictors (e.g., anxiety and depressive symptoms), this shared component tends to be quite large. Partialling it out often leads to an

overestimation of the role played by one component, and to an underestimation of the role played by the other components. A similar issue occurs when considering students' ratings of behavioral, emotional, and cognitive engagements, which also tends to be highly correlated with one another, and yet to each have their own unique character (Wang et al., 2019).

A bifactor approach makes it possible to identify a global factor (G-factor) together with subscale-specific orthogonal factors (S-factors). The G-factor reflects the commonality shared among all items (e.g., global levels of internalizing behaviors across dimensions). The S-factors are orthogonal (e.g., specific levels anxiety and depressive symptoms) and reflect the variance shared across items forming each subscale beyond that already explained by the G-factor (Morin, Boudrias, Marsh, Madore, et al., 2016). In plain language, this approach separates ratings into independent (i.e., uncorrelated) components. Thus, this approach simultaneously considers the role of students' global levels of internalizing or externalizing behaviors, together with the unique predictive value of anxiety, depressive symptoms, hyperactivity/inattention, and opposition/defiance beyond these global levels. The ability of this approach to achieve a more accurate representation of behavior problems (e.g., Caci et al., 2015; Caspi et al., 2014) and engagement (e.g., Wang et al., 2019; Wang M.T. et al., 2016) has been previously demonstrated. However, this approach has yet to be implemented in research focusing on the nexus between behavior problems and engagement (see Table 1).

Current Study

Scholars and school practitioners all agree that students' externalizing and internalizing behavior problems cause a threat to their active engagement in school. However, it remains unclear if one type of behavior problem is a more important risk factor for student engagement and how these risks may differ between boys and girls and between different developmental periods. This research assesses the relations between students' externalizing (global, hyperactivity/inattention, and opposition/defiance) and internalizing (global, anxiety, and depressive symptoms) behaviors and their global and specific levels of engagement (behavioral, emotional, and cognitive). This study also investigates whether these associations are similar or differ across samples of elementary (Study 1; grades 3 to 6) and secondary (Study 2; grades 7 and 8) students, as well as between boys and girls. Studying these associations before and after the transition to secondary school can inform about student adjustment during this period, where they are known to be particularly at risk of disengagement (Eccles & Roeser, 2009). More specifically, the present research pursues two objectives, via two independent studies:

The first objective of this study is to investigate the associations between global and specific levels of externalizing and internalizing behaviors and student engagement. This study anticipates that both types of behavioral problems will share negative associations with student engagement levels, both in elementary and secondary school students. Yet, existing results are mixed regarding which of these problems has the strongest association with which dimension of engagement. As such, the differential effects of global and specific levels of internalizing and externalizing behaviors in the prediction of global relative to specific components of student engagement is left as an open research question. Moreover, although different developmental trends have been reported between boys and girls, most studies that have assessed sex differences in the nexus linking behavior problems to student engagement have found no such differences (see Table 1). Thus, the study anticipates that these associations will also be similar for boys and girls.

The second objective of this study is to verify whether there are interactive effects between behavior problems in predicting student engagement. Finding a significant interaction would be consistent with the presence of a risk accumulation effect due to the co-occurrence of various behavior problems. More specifically, the study assesses the interactions between high global levels of externalizing and internalizing behaviors; between high global levels of externalizing behaviors and high specific levels of internalizing behaviors (anxiety, and depressive symptoms); between high global levels of internalizing behaviors and high specific levels of externalizing behaviors (hyperactivity/inattention, and opposition/defiance); and between high specific levels of externalizing and internalizing behaviors. Although the co-occurrence of behavior problems is reported in several studies, it is unclear if this risk accumulation causes extra threats to student engagement. Still, co-occurring externalizing and internalizing behaviors seem more problematic when they appear early in development. Thus, the study hypothesizes that this risk accumulation (i.e., combining multiple types of behavioral problems) might be associated with a further reduction in engagement, particularly in the elementary school sample.

Study 1 (Elementary School Sample)

Methods

Sample and procedure. Study 1 relies on a sample of 1,036 3rd to 6th grade students recruited in seven elementary schools from one school board located in the Canadian province of Quebec. The majority of students were Caucasian and came from middle-class families, which is representative of the student population outside of the Montreal area (MEES, 2019). Students were, on average, 9.94 years old (*SD* = 1.30), and 47.30% of them were girls. This project was approved by the University's research ethics committee. Prior to data collection, the research team obtained active parental, student, and teacher consent. In November 2011, students answered a 45 minutes computerized questionnaire on their school experiences, including all measures used in the present study. During data collection, each classroom was supervised by two trained research assistants, and teachers used this

time to complete a paper questionnaire on the behaviors and performance of each of their students.

Measures[2]. *Externalizing behaviors.* Teachers rated each student using two scales from the French adaptation (Capron et al., 2007) of the Strengths and Difficulties Questionnaire (Goodman, 2001): Hyperactivity/Inattention (five items; α = .764; e.g., "This child is restless, overactive, cannot stay still for long") and Oppositional/Defiant (four items[3]; α = .725; e.g., "This child often loses temper"). Each item was rated on a three-point scale (1- *not true* to 3-*certainly true*).

Internalizing behaviors. Student rated their own levels of internalizing behaviors using two scales from the French adaptation (Trembaly et al., 1987) of the Preschool Behavior Questionnaire (Hoge et al., 1985): Anxiety (four items; α = .743; e.g., "You worry that you are not as good as other students") and Depressive symptoms (three items; α = .779; e.g., "You are unhappy or sad"). Each item was rated on a three-point response scale (1- *not true* to 3-*certainly true*).

Student engagement. Students completed the Dimensions of School Engagement Scale (Archambault & Vandenbossche-Makombo, 2014) in relation to their language class (i.e., French). The behavioral (α = .685; e.g., "I follow my teachers' instructions"), emotional (α = .770; e.g., "I think that reading and writing assignments are interesting"), and cognitive (α = .737; e.g., "When I finish an assignment, I check to make sure that I did not make mistakes") engagement subscales each comprised three items rated on a five-point scale (1- *not at all* to 5- *very much*).

Covariates. Students self-reported their sex (0 = male; 1 = female), grade level (3rd, 4th, 5th, or 6th grade), and family status (0 = nuclear family; 1 = other). Teachers reported the achievement of their students in language class (i.e., French) by comparing each of them to the class average on a scale ranging from (1) *significantly below average* to (5) *significantly above average* (Duncan et al., 2007).

Analyses. All analyses were performed with Mplus 8.2 (Muthén & Muthén, 2019). Alternative measurement models were first estimated and contrasted to establish the optimal factor solution (CFA or bifactor-CFA) for each set of variables (externalizing behaviors, internalizing behaviors, and student engagement) (see Appendix 1 for details). The selection of the optimal and most invariant measurement model for all constructs was followed by the estimation of predictive models designed to assess the role of externalizing and internalizing behaviors in the prediction of student engagement. Given the complexity of the measurement models estimated in this study, it was not possible to estimate these predictive models using a fully latent approach. Rather, the study relied on factor scores saved from the optimal measurement model identified previously. This allowed relying on the Maximum Likelihood Robust (MLR) estimator for these models, together with Full Information Maximum Likelihood procedures (Enders, 2010) to handle missing data (5.77% to 28.10%) on the covariates.

A first predictive model including only the covariates (sex, grade, family status, and achievement) and their association with student engagement was then estimated. To maximize model parsimony, only the covariates associated in a statistically significant manner with each outcome were retained in the main analyses in an outcome-specific manner. In a second model, externalizing and internalizing behaviors were added as additional predictors. In a third model, two-way interactions between the externalizing and internalizing behaviors factors were also included in order to address Objective 2. Statistically significant interaction effects were interpreted by the examination of simple slopes depicting the effect of the predictor at different levels (-1*SD*, *M*, and +1*SD*) of the moderator (Marsh et al., 2013). Finally, the predictive equivalence of these relations across sexes was assessed through the following sequence of models: (i) configural equivalence (same predictive model with no additional constraint); (ii) equivalence of the regression slopes; (ii) equivalence of the regression intercepts; (iii) equivalence of the regression residuals.

Model fit was assessed using the chi-square statistic (χ^2), the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), and the Tucker-Lewis Index (TLI) (Marsh et al., 2005). RMSEA values smaller than 0.08 and 0.06 respectively suggest acceptable and excellent model fit. Values above .90 and .95 for the CFI and TLI respectively indicate adequate and excellent model fit. In tests of measurement invariance and predictive equivalence, increases in RMSEA of more than .015 and decreases in CFI and TLI of more than .010 were considered to indicate non-invariance (Chen, 2007).

Results

Correlations between the various factor scores and covariates are reported in Table 2. Results from the model including only covariates are reported on the top section of Table 3. Based on this model, family structure was removed from further analyses, and only significant paths involving student sex, grade level, and achievement were kept in the following models (sex and grade level in the prediction of global engagement and specific emotional engagement, and achievement in the prediction of global engagement, and specific behavioral and emotional engagement). The main predictive model had an excellent level of fit to the data (χ^2 = 1.453, df = 5, *p*

[2] See Appendix 2 for a detailed description of the measurement models.

[3] The original scale included a fifth item ("This student steals from home, school, or elsewhere."). However, because of the rarity of this behavior in elementary school children and the difficulty for teachers to observe "hidden" behaviors, this item was not included in this study.

= .918; RMSEA = .000; CFI = 1.000; TLI = 1.042). These results (see Table 3) indicate that the Externalizing behaviors G-factor was negatively associated with the Engagement G-factor and Behavioral engagement S-factor. The Internalizing behaviors G-factor was also negatively associated with the Engagement G-factor and Cognitive engagement S-factor. Moreover, the Anxiety S-factor was associated with lower levels on the Engagement G-factor and the Emotional engagement S-factor. The Hyperactivity/Inattention, Oppositional/Defiant, and Depressive symptoms S-factors did not contribute to the prediction of any engagement factor. Adding interactions between G-factors, between G-factors and S-factors, and between S-factors did not further contribute to explaining any of the engagement G- and S-factors. Tests of predictive equivalence across sexes supported the equivalence of the regression slopes, but suggest sex differences related to the regression intercepts associated with the Engagement G-factor and the Emotional engagement S-factor. More precisely, these regression intercepts were respectively .449 and .466 SD higher among girls relative to boys, suggesting that girls tend to present higher levels (roughly .5 SD) on these S-factors relative to boys presenting matching levels of internalizing and externalizing behaviors. Likewise, the results also revealed sex differences related to the regression residual of the Cognitive engagement S-factor, showing that the model was able to explain 2.2% of the variance of Cognitive engagement for boys compared to 3.0% for girls.

Study 2 (Secondary School Sample)

Methods

Sample and procedure. Study 2 relies on a sample of 1,011 7th and 8th grade secondary school students recruited in four schools located in the Canadian province of Quebec. Although collected in the same Canadian province, this sample is independent from, and unrelated to, the sample used in Study 1. Students were on average 12.93 years old (s.d. = 0.76), 55.77% of them were girls, 26.50% were first-generation immigrants (born abroad), and 34.14% were second-generation immigrants (born in Canada, with one or both parents born abroad), and 39.26% were third-generation plus. This project was approved by the University's research ethics committee. Active written parental consent was obtained for all students. Students also gave their active consent to participate. In Fall 2018, students answered a 45 minutes computerized questionnaire on their well-being and school experiences including all measures used in the present study. During data collection, each classroom was supervised by a trained research assistant.

Measures[4]. ***Externalizing behaviors.*** Student self-reported their externalizing behaviors using two scales from the French adaptation (Capron et al., 2007) of the Strengths and Difficulties Questionnaire (Goodman, 2001): Hyperactivity/Inattention (five items; α = .654; e.g., "I am restless, I cannot stay still for long") and Oppositional/Defiant (five items; α =.582; e.g., "I lose my temper easily or often get angry"). Each item was rated on a three-point scale ranging from (1) *not true* to (3) *certainly true*.

Internalizing behaviors. Student self-reported their internalizing behaviors using two scales. Anxiety was assessed using nine items (α = .864; e.g., "You worry that you are not as good as other students") from the Screen for Child Anxiety Related Emotional Disorders (Birmaher et al., 1999) validated in French (Blais et al., 2002). These items were rated using a three-point scale ranging from (1) *almost never* to (3) *often*. Depressive symptoms were measured using six items (α = .932; e.g., "You feel sad") from a short French version (Blais et al., 2002) of the Center for Epidemiologic Studies Depression scale (Radloff, 1977). These items were rated on a six-point scale ranging from (1) *never* to (6) *always*.

Student engagement. Students completed a measure of engagement directly developed and validated in French (Archambault, Janosz, Fallu et al., 2009). This scale is not specific to any school subject and covers three dimensions: Behavioral (four items; α = .712; e.g., "I have disrupted the classroom on purpose"), emotional (six items; α = .851; e.g., "I enjoy what we do in school"), and cognitive (seven items; α = .895; e.g., "I take time to make sure that I understand assignments"). Items were rated on a five-point response scale ranging from (1) *strongly disagree* to (5) *strongly agree*.

Covariates. Students self-reported their sex (0 = male; 1 = female), grade level (7th or 8th grade), family status (0 = nuclear family; 1 = other), grade retention (0 = never; 1 = repeated a grade at least once), and immigration status (0 = non-immigrant; 1 = 1st and 2nd generation immigrants).

Analyses. The same procedure as in Study 1 was applied to Study 2. Given that Study 2 was conducted among a multiethnic sample, measurement invariance and predictive equivalence were also tested between groups of immigrant (1st and 2nd generation) and non-immigrant students (3rd generation+). All measures were student-reported using a computerized method allowing for no missing data.

Results

Correlations between factors and covariates are reported in Table 2. Results from the model including only covariates are reported in the bottom section of Table 3. Based on this model, only significant paths involving student sex, grade level, retention, family structure, and immigration status were kept in the following models (sex in the prediction of all outcomes, grade level in the prediction of global engagement and specific behavioral and emotional engagement, family status in the prediction of global engagement and specific emotional

[4] See Appendix 2 for a detailed description of the measurement models.

engagement, retention in the prediction of specific behavioral engagement, and immigration status in the prediction of all three specific engagement dimensions). The predictive model had an excellent level of fit to the data (χ^2 = 9.889, df = 7, p = .195; RMSEA = .020; CFI =.998; TLI =.984). These results (see Table 3) indicate that the Externalizing behaviors G-factor and Hyperactivity/Inattention S-factor were negatively associated with the Engagement G-factor and with the Behavioral engagement S-factor. The Oppositional/Defiant S-factor was also negatively associated with the Behavioral engagement S-factor. The Internalizing behaviors G-factor and the Anxiety S-factor were not directly associated with any dimension of engagement. The Depressive symptoms S-factor negatively contributed to the Engagement G-factor and to the Emotional engagement S-factor. Finally, no variable contributed to the prediction of the Cognitive Engagement S-factor.

When adding interactions between G-factors, between G-factors and S-factors, and between S-factors, model fit remained excellent (χ^2 = 10.460, df = 7, p = .164; RMSEA = .022; CFI =.997; TLI =.972) and the significance of all direct effects remained unchanged. These results revealed two statistically significant interactions. Thus, the interaction between the Internalizing behaviors G-factor and the Hyperactivity/Inattention S-factor was significantly associated with the Engagement G-factor (β = .086, p = .013) and with the Emotional engagement S-factor (β = .091, p = .009). The results from simple slopes analyses are graphically presented in Figures 1 and 2. These results indicate that, for students with a low level on the Hyperactivity/Inattention S-factor, the Internalizing behaviors G-factor was associated with a lower level on the Engagement G-factor and the Emotional engagement S-factor. For students with a high level on the Hyperactivity/Inattention S-factor, the Internalizing behaviors G-factor was associated with a higher level on these two dimensions.

Tests of predictive equivalence across immigration status supported the equivalence of the regression slopes and regression residuals, but not of the regression intercepts. More precisely, immigrants tended to present lower intercepts on the Behavioral engagement S-factor (-.233 SD) relative to non-immigrants presenting similar levels of internalizing and externalizing behaviors. Tests of predictive equivalence also revealed sex differences related to some regression slopes and intercepts but supported the equivalence of the regression residuals. In fact, a single regression slope, related to the association between immigration status (a covariate) and the Emotional engagement S-factor was found to differ as a function of sex, being statistically significant for girls (b = –.197, s.e. = .057, p < .001, β = –.127; showing that non-immigrant girls had higher specific levels of emotional engagement than immigrant girls), but not for boys (b = .023, s.e. = .060, p = .701, β =.014). In terms of regression intercepts, girls presented higher scores on the Behavioral (.140 SD) and Cognitive (.125 SD) engagement S-factor relative to boys presenting similar levels of behavioral problems.

Discussion

Students with externalizing or internalizing behavior problems are at increased risk of not completing upper secondary school (MEES, 2019), partly due to their lower engagement (Gut et al., 2012). Considering that these students represent an important proportion of those enrolled in regular schools (Olivier et al., 2018), researchers and practitioners strive to identify priority prevention and intervention targets to minimize their risk of disengagement. Behavior problems can be conceptualized as global (externalizing or internalizing) and specific (hyperactivity/inattention, opposition/defiance, anxiety, and depression) difficulties, which are likely to play a distinct role on student disengagement (Caci et al., 2015). However, it is still unclear which type of behavior problem, internalizing or externalizing, are associated with higher risk of disengagement for boys and girls at different developmental periods. Whether and how the accumulation of different behavior problems could pose an additional risk for student school functioning also remain unknown (Eisenberg et al., 2009). To identify more specific intervention targets, this study first aimed to disentangle the global and specific contribution of externalizing and internalizing behavior problems on the behavioral, emotional, and cognitive engagement of elementary and secondary school students. Second, the study assessed the association between global and specific behavior problems with student engagement, while also contrasting the results between boys and girls.

More specifically, this study sought to investigate the combined role played by global and specific levels of externalizing and internalizing behaviors in the prediction of students' global and specific levels of school engagement during the end of elementary school (Study 1) and the beginning of secondary school (Study 2). The study also considers the possible effects of sex (Study 1 and 2) and immigration status (Study 2 only). The results show that the global aspects of student externalizing and internalizing problems are as important as their specific facets to understand threats to student engagement. Although not longitudinal in nature, the juxtaposition of results obtained from these two studies suggests that developmental trends might influence the nature and strength of the associations between students' behavior adjustment problems and their levels of school engagement.

Externalizing and internalizing behaviors seem to function slightly differently according to the developmental period. First, externalizing behaviors appeared important risk factors for students' global engagement levels and their specific levels of behavioral engagement across the elementary and secondary school years. These effects appeared not to be limited to students' global levels of externalizing behaviors during the elementary school years, but also encompass their specific hyperactivity/inattention and opposition/defiance during the secondary school years. Second, the role of internalizing behaviors changed more importantly across these two developmental periods. Thus, students' global levels of internalizing behaviors and their specific levels

of anxiety were found to be associated with lower global levels of engagement during the elementary school years. During this period, students displaying high anxiety, beyond their global internalizing behaviors, also reported lower specific levels of emotional engagement. In contrast, during the secondary school years, only students displaying high specific levels of depressive symptoms, beyond their global level of internalizing behaviors, reported lower global levels of school engagement and lower specific levels of emotional engagement.

Overall, the results suggest that at least some of the global and specific facets of behavior adjustment problems appear to play a role reducing some aspect of student engagement over the course of youth development. However, the mechanisms underpinning these effects might be submitted to developmental differences. Moreover, beyond their negative association with students' global levels of school engagement, externalizing behavior problems also shared negative associations with students' specific levels of behavioral engagement, whereas internalizing behaviors also shared negative associations with students' specific levels of emotional engagement. These specific associations are consistent with the view that externalizing behaviors may prevent the optimal behavioral investment of students because of self-regulation difficulties (e.g., Campbell et al., 2014). Internalizing behaviors may rather impede students' positive emotional investment because of cognitive biases or emotional dysregulation, which may alter their cognitive availability (Garber & Rao, 2014).

Externalizing Behaviors and Student Engagement

As expected, externalizing behaviors were found to be associated with lower levels of student engagement in a way that was highly similar across the elementary and secondary school samples. More precisely, global levels of externalizing behaviors were associated with lower global levels of school engagement and lower specific levels of behavioral engagement in both samples, as well as for boys and girls. Thus, irrespective of developmental stage and sex, externalizing behaviors seem to remain incompatible with the behavioral demands of a classroom, as well as with displaying a globally engaged attitude toward school.

Beyond the effects of students' global levels of externalizing behaviors, their specific levels of hyperactivity/inattention and oppositional-defiance were found to further contribute to reducing their global levels of engagement as well as their specific levels of behavioral engagement in the secondary school sample. There are a few possible explanations for these observations. First, as students enter secondary school, they are exposed to more diverse and complex subjects but also have to navigate a new environment, while receiving less intensive support from teachers and school professionals (Eccles & Roeser, 2009). In this context, high specific levels of hyperactivity/inattention and opposition/defiance, and the related self-regulation impairments, may become additional threats to their full engagement in school. From a prevention perspective, these results suggest that targeting youth displaying very high specific levels of hyperactivity/inattention or opposition/defiance, as it is often done in various school systems, may prove fruitful. However, these results also indicate that students displaying a wider range of externalizing behaviors may benefit from preventive interventions, even if their levels of behavioral problems do not appear to be as extreme as that of their peers in specific areas.

Second, the source of information used to obtain ratings of externalizing behaviors in the present study may play a role in these differential findings. In the secondary school sample, students self-reported their own externalizing behaviors, whereas in the elementary school sample, teachers rated the externalizing behaviors of their students. Teachers are known to rely on a more global or holistic assessment process when asked to describe their students' behaviors, which might have made it harder to detect effects associated with more specific types of behaviors in this sample (Flake & Petway, 2019). In contrast, secondary school students were asked to rate both their externalizing behaviors and their school engagement, which might have slightly inflated the associations between these two types of constructs.

Another noteworthy difference is related to the items used to measure student behavioral engagement. In the elementary school sample, these items mainly focused on observable deviations from complying with classroom expectations. In contrast, in the secondary school sample, these items focused on more widespread types of disruptive behaviors. This more generic coverage may make it easier to differentiate the predictive role of global levels of externalizing behaviors from that of behaviors more specifically related to hyperactivity/inattention and opposition/defiance. In contrast, it might be harder to break down the effects of classroom compliance into specific dimensions of externalizing behavior. Future research is needed to assess the plausibility of these alternative explanations more systematically. However, given the consistency of results obtained across samples and sex, these results provide substantial evidence that global externalizing behaviors seem to represent threats to student engagement both before and after the transition to secondary school.

Internalizing Behaviors and Student Engagement

As expected, internalizing behaviors also appeared to threaten student engagement. However, the mechanisms underpinning these effects seem to change over the course of development. In the elementary school sample, students' global levels of internalizing behaviors were found to be associated with lower global levels of engagement and with lower specific levels of cognitive engagement. This result indicates that the combination of anxiety and depressive symptoms seems to prevent students from being fully engaged in their learning experiences. Moreover, students' global level of internalizing behaviors also prevented their full cognitive engagement. Although unexpected, this result is consistent with the view that students' emotional dysregulation

may lead to difficulties in self-regulated and deep processing learning strategies.

However, the size of these effects was not as large as those associated with global levels of externalizing behaviors, even though teachers' reported students' externalizing behaviors, whereas students self-reported their levels of internalizing behaviors and engagement in this sample. Effect size disparities were also reported in earlier studies (Searle et al., 2013), suggesting that the stronger effects of externalizing, relative to internalizing, behaviors on student engagement may not be specific to the present study. In addition, specific levels of anxiety were also found to be related to lower levels of global engagement and with lower specific levels of emotional engagement among elementary students. This suggests that in younger students, anxiety may play a specific role in impairing student engagement, especially when the emotional aspects of student engagement are considered (Kurdi & Archambault, 2020). The higher prevalence of anxiety than depressive symptoms in younger children (Vasey et al., 2014) might explain why high specific levels of anxiety appear to be significant threats to global levels of engagement during this developmental period. Indeed, students might not yet have learned to appropriately self-regulate their anxious thoughts and emotions.

In the secondary school sample, only specific levels of depressive symptoms were found to be directly associated with students' lower global levels engagement and with lower specific levels of emotional engagement. These two dimensions are the same that were found to be affected by specific levels of anxiety in the elementary school sample. Contrasting these results suggests a developmental difference, which may be related to the greater prevalence of anxiety (relative to depressive symptoms) in younger samples (Vasey et al., 2014), and to the increasing rates of depressive symptoms observed among adolescents (Garber & Rao, 2014). Students displaying anxiety or a combination of anxiety and depressive symptoms (i.e., global levels of internalizing behaviors) at a younger age may have progressively learned to control and self-regulate their emotions and thoughts in order to protect their ability to stay engaged at school despite these negative emotions. This would be detected as a lack of associations between these types of internalizing behaviors and engagement in the secondary school sample. Such results are consistent with those previously reported by Garvik et al. (2014), demonstrating only small associations between internalizing behaviors and student engagement among a sample of adolescents. In contrast, students' experiencing a rise in depressive symptoms during adolescence may not have learned to control the impact of these symptoms as their peers with an earlier onset of internalizing behaviors. This would then explain that the negative association found for anxiety in younger children shifts to depressive thoughts in young adolescents, who have yet to learn to manage these negative feelings.

Accumulation of Behavioral Risks

The results did not uncover the expected effect of an accumulation of behavioral risks in the prediction of engagement levels in any of the samples or subsamples considered in the present study. Moreover, the interaction effects found within the secondary school sample showed that for students presenting high levels of hyperactivity/inattention, global levels of internalizing behaviors were associated with higher global levels of engagement and with higher specific levels of emotional engagement. In contrast, for students displaying low levels of hyperactivity/inattention, global levels of internalizing behaviors were associated with lower global levels of engagement, and with lower specific levels of emotional engagement.

This unexpected interaction could indicate two possible phenomena. First, it is possible that these two behavioral risk factors, when accumulated, tap into incompatible processes. As such, having a tendency to be very active and distracted while simultaneously being concerned, nervous, and sad might exert counterbalancing effects, allowing students to maintain adequate engagement levels. Being very agitated could be a strategy for these students to deal with anxious and depressive thoughts. In a review of the literature, Jarrett et al. (2014) have shown that internalizing behaviors tend to lead to lower levels of externalizing behaviors (also see Morin et al., 2017). These observations thus suggest at least some level of incompatibility between externalizing and internalizing behaviors. Second, it is possible that secondary school students characterized by multiple types of behavioral problems might have been targeted for school-based interventions or provided access to extra support mechanisms to maintain a satisfactory level of engagement. A study by Olivier et al. (2018) did find that boys who presented a combination of externalizing and internalizing behaviors were less at risk of disengagement than students displaying only one type of such behaviors. In any case, future research is needed to investigate the plausibility and complementarity of these two explanatory hypotheses.

Universality of Behavior Problems as Risk Factors

The associations between students' externalizing and internalizing behavior problems and engagement were found to be the same between boys and girls, and between students from immigrant and non-immigrant backgrounds. This suggests that, although the levels of behavior problems and engagement may vary as a function of sex or immigration status (see Appendix 2 and 3), externalizing and internalizing behavior problems are likely universal risk factors. As such, these risks should not be subject to differential treatment from school professionals. Similarly, some studies found that behavior problems influence boys and girls equally (opposition/defiance: Archambault et al., 2017; hyperactivity/inattention: Demaray & Jenkins, 2011; depressive symptoms: Garvik et al., 2014; anxiety: Kurdi & Archambault, 2020) as well as immigrants and non-immigrants (aggression: Konold et al., 2016; hyperactivity/inattention: Tardif-Grenier et al., 2019). Yet, these results are far from consensual as

others found boys and girls (anxiety and depressive symptoms: Derdikman-Eiron et al., 2011; externalizing and internalizing: Olivier et al., 2018) and youths from different cultures (anxiety: Gillen-O'Neel et al., 2011) to be differently impacted, especially by internalizing behaviors. Globally, these results call for a more systematic investigation of the potential moderating roles of sex and immigration background.

Limitations

The first limitation is that none of the two studies relied on a longitudinal design. As such, the hypothetical developmental trends identified in this study could also be the result of a cohort effect (Little, 2013) and that reciprocal effects of engagement on behavior problems might also be at play. Moreover, engagement tends to remain relatively stable over time both in childhood (Archambault & Dupéré, 2016) and adolescence (Wang & Eccles, 2012). Yet, it was not possible to control for preexisting levels of engagement, which would have strengthened the results.

Second, because this investigation combines two independent samples of participants initially collected for different purposes, the study had to rely on measures that were not exactly equivalent in the two samples despite being validated for each developmental period. As typically done in studies assessing externalizing behaviors and student engagement (see Table 1), teachers' ratings of students' externalizing behaviors were obtained in the elementary school sample, and students' self-reports of these same behaviors were used in the secondary school sample. This decision is consistent with the greater ease with which elementary teachers can report on the externalizing behaviors of children with whom they interact all day, five days a week (Smith, 2007). In contrast, secondary school teachers only have limited contact with their students, typically limited to a few hours of in-class teaching, making it virtually impossible for them to report reliably on the whole range of externalizing behaviors exhibited by their students (Randazzo et al., 2003). Conversely, the measure of internalizing behaviors was self-reported in both samples. Given the naturally internal and hard-to-observed nature of these behaviors, youth are considered more reliable informants (Smith, 2007). Also, the measure was longer in the secondary school sample than in the elementary school sample. Respectively, each measure was selected to reflect best students' developmental stage, as well as the normative attention span of youth from this age group (which is typically longer in adolescents than in children; Fuchs, 2005). Studies have shown that different questionnaires measuring internalizing behaviors can function equally well even if one is more extensive than the other (Dart et al., 2020).

The student engagement measure was also not equivalent across samples, as the two measures used have only been validated in their respective age group (elementary or secondary school samples). The elementary school measure focused on language class, whereas the secondary school measure encompassed students' more general school engagement. Although the emotional dimension covered relatively similar emotions but did so more extensively in the secondary school sample, the behavioral dimension had a stronger focus on compliance among elementary students, and disruptiveness among secondary students. Conversely, the cognitive dimension was more extensively covered in secondary students, whose self-regulation abilities tend to be more developed than those of elementary students. These two dimensions are thus consistent with known developmental differences in relation to student engagement between younger (Archambault & Dupéré, 2016) and older students (Wang & Eccles, 2012). As for the internalizing behavior measure, the longer length of the secondary school measure of engagement aimed to match adolescents' greater attention span. However, even if all measures were selected to best represent the developmental stage of the students, it remains that the reliance on distinct measures may complicate the comparison of results across studies, especially with the specific factors. Thus, this limitation calls for replication with measures that are perhaps more comparable across developmental periods. Still, these differences in measures also reinforce the robustness of the results that are common to both studies.

Third, considering the number of parameters included in the models, the complexity of bifactor models, and the sample size, it was impossible to rely on a fully latent approach. To overcome this limitation, the study relied on factor scores, which still control for a part of measurement error (DiStefano et al., 2009). Finally, neither this investigation nor the two studies that provided the data used in the present investigation aimed to assess ethnic or immigration status differences. For this reason, one sample was relatively homogenous (elementary school), whereas the other only provided limited information about the students' immigration background (secondary school). This information allowed conducting tests of measurement and predictive invariance as a function of immigration status among secondary school students and controlling for this variable as part of the main analyses of this study. These verifications revealed very few differences, consistent with the robustness of the findings across distinct immigration groups. Still, due to the very limited amount of information available in this regard, this study cannot be considered to represent a systematic assessment of immigration differences, and additional studies relying on more diverse samples of children and adolescents will be required to understand better the role played by culture, ethnicity, and nationality in these relations.

Conclusion

This study formally assessed the joint risk posed by global and specific levels of externalizing and internalizing behavior problems for student engagement. The negative role of externalizing behavior on students' global levels of engagement and their specific levels of behavioral engagement remained salient in elementary

and secondary school students. In contrast, age differences were found pertaining to internalizing behaviors. Among elementary school students, specific levels of anxiety, and global levels of internalizing behaviors were important risk factors for reduced engagement (global, emotional, and cognitive). Among secondary school students, specific levels of depressive symptoms became more prominent risk factors for global and emotional engagement. Despite these possible developmental effects, the results highlight that externalizing and internalizing behaviors represent risk factors for school disengagement equally among boys and girls, as well as immigrants and non-immigrants. Moreover, the results do not support that students who accumulate behavior problems are more at risk of disengagement than the risk already posed by displaying any one type of problem. These results generally suggest that prevention strategies targeting externalizing behaviors should be favored before and after the transition to secondary school. In contrast, prevention strategies designed to minimize the impact of internalizing behaviors could focus on anxiety and global levels of internalizing behaviors among elementary students and on depressive symptoms following the transition to secondary school.

References

Achenbach, T.M. & Edelbrock, C.S. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin, 85*(6), 1275-1301.

Al-Hendawi, M. (2012). Academic engagement of students with emotional and behavioral disorders: Existing research, issues, and future directions. *Emotional and Behavioural Difficulties*, *17*(2), 125-141.

Archambault, I., & Dupéré, V. (2016). Joint trajectories of behavioral, affective, and cognitive engagement in elementary school. *The Journal of Educational Research, 110*(2), 188-198.

Archambault, I., Janosz, M., Fallu, J.S., & Pagani, L.S. (2009). Student engagement and its relationship with early high school dropout. *Journal of Adolescence, 32*(3), 651-670.

Archambault, I., Janosz, M., Morizot, J., & Pagani, L.S. (2009). Adolescent behavioral, affective, and cognitive engagement in school: Relationship to dropout. *Journal of School Health, 79*(9), 408-415.

Archambault, I., & Vandenbossche-Makombo, J. (2014). Validation de l'échelle des dimensions de l'engagement scolaire (ÉDES) chez les élèves du primaire. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement*, *46*(2), 275-288.

Archambault, I., Vandenbossche-Makombo, J., & Fraser, S.L. (2017). Students' oppositional behaviors and engagement in school: The differential role of the student-teacher relationship. *Journal of Child and Family Studies, 26*(6), 1702-1712.

Baker, J.A., Grant, S., & Morlock, L. (2008). The teacher-student relationship as a developmental context for children with internalizing or externalizing behaviors. *School Psychology Quarterly, 23*(1), 3-15.

Bierman, K.L., & Sasser, T.R. (2014). Conduct disorder. In M. Lewis & K.D. Rudolph (Eds.), *Handbook of Developmental Psychopathology, 3rd Ed.* (pp. 467-487). Springer.

Birmaher, B., Brent, D.A., Chiappetta, L., Bridge, J., Monga, S., & Baugher, M. (1999). Psychometric properties of the Screen fo Child Anxiety Related Emotional Disorders (SCARED): A replication study. *Journal of the American Academic Child and Adolescent Psychiatry*, *38*(10), 1230-1236.

Blais, M., Hess, U., & Riddle, A.S. (2002). A multi-group investigation of the CES-D's measurement structure accrcoss adolesccents, young adults and middle-aged adults. *CIRANO Working Papers, 2002s-36.*

Caci, H., Morin, A.J., & Tran, A. (2015). Investigation of a bifactor model of the Strengths and Difficulties Questionnaire. *European Child and Adolescent Psychiatry, 24*(10), 1291-1301.

Campbell, S.B., Halperin, J.M., & Sonuga-Barke, E.J.S. (2014). A developmental perspective on Attention-Deficit/Hyperactivity Disorder (ADHD). In M. Lewis & K.D. Rudolph (Eds.), *Handbook of Developmental Psychopathology, 3rd Ed.* (pp. 427-448). Springer.

Capron, C., Thérond, C., & Duyme, M. (2007). Psychometric properties of the French version of the self-report and teacher Strenghts and Difficulties Questionnaire (SDQ). *European Journal of Psychological Assessment, 23*(2), 79-88.

Caspi, A., Houts, R.M., Belsky, D.W., Goldman-Mellor, S.J., Harrington, H., Israel, S., Meirer, M.H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T.E. (2014). The p Factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science, 2*(2), 119-137.

Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464-504.

Curhan, A.L., Rabinowitz, J.A., Pas, E.T., & Bradshaw, C.P. (2020). Informant discrepancies in internalizing and externalizing symptoms in an at-risk sample: The role of parenting and school engagement. *Journal of Youth and Adolescence, 49*(1), 311-322.

Dart, E.H., Arora, P., Collins, T.A., Stark, K., Cook, C.R., Duong, M.T., McCarthy, C.A., & Doll, B. (2020). Developing a brief behavior rating scale for progress monitoring of depression in school settings. *Assessment for Effective Intervention*, 45(3), 163-172.

Demaray, M.K., & Jenkins, L.N. (2011). Relations among academic enablers and academic achievement in children with and without high levels of parent-rated symptoms of inattention, impulsivity, and hyperactivity. *Psychology in the Schools, 48*(6), 573-586.

Derdikman-Eiron, R., Indredavik, M.S., Bratberg, G.H., Taraldsen, G., Bakken, I.J., & Colton, M. (2011). Gender differences in subjective well-being, self-esteem and psychosocial functioning in adolescents with symptoms of anxiety and depression: Findings from the Nord-Trondelag Health Study. *Scandinavian Journal of Psychology, 52*(3), 261-267.

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation, 14*(20), 1-11.

Dorio, N. B., Secord Fredrick, S., & Demaray, M. K. (2019). School engagement and the role of peer victimization, depressive symptoms, and rumination. *Journal of Early Adolescence, 39*(7), 962-992.

Duncan, G.J., Dowsett, C.J., Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and achievement. *Developmental Psychology, 43*(6), 1428-1446.

Eccles, J.S. & Roeser, R.W. (2009). Schools, academic motivation, and stage-environment fit. In R.M. Lerner &

L. Steinberg (Eds.). *Handbook of Adolescent Psychology*, (pp. 404-434). Wiley.
Eisenberg, N., Valiente, C., Spinrad, T.L., Liew, J., Zhou, Q., Losoya, S.H., Reiser, M. & Cumberland, A. (2009). Longitudinal relations of children's effortful control, impulsivity, and negative emotionality to their externalizing, internalizing, and co-occurring behavior problems. *Developmental Psychology, 45*(4), 988-1008.
Enders, C. K. (2010). *Applied missing data analysis*. Guildford.
Evans, G. W., Li, D., & Whipple, S.S. (2013). Cumulative risk and child development. *Psychological Bulletin, 139*(6), 1342-1396.
Fiorilli, C., De Stasio, S., Di Chiacchio, C., Pepe, A., & Salmela-Aro, K. (2017). School burnout, depressive symptoms and engagement: Their combined effect on student achievement. *International Journal of Educational Research, 84*, 1-12.
Flake, J.K., & Petway, K.R. (2019). Methodologies for investigating and interpreting student-teacher rating incongruence in noncognitive assessment. *Educational Measurement: Issues and Practice, 38*(1), 63-77.
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.
Fredricks, J.A., Filsecker, M., & Lawson, M. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and method issues. *Learning and Instruction, 43*, 1-4.
Fuchs, M. (2005). Children and adolescents as respondents. *Journal of Official Statistics*, *21*, 701-725
Garber, J., & Rao, U. (2014). Depression in children and adolescents. In M. Lewis & K.D. Rudolph (Eds.), *Handbook of Developmental Psychopathology, Third Edition* (pp. 489-520). Springer.
Garvik, M., Idsoe, T., & Bru, E. (2014). Depression and school engagement among norwegian secondary vocational school students. *Scandinavian Journal of Educational Research, 58*(5), 592-608.
Gillen-O'Neel, C. G., Ruble, D. N., & Fuligni, A. J. (2011). Ethnic stigma, academic anxiety, and intrinsic motivation in middle childhood. *Child Developement, 82*(5), 1470-1485.
Gogol, K., Brunner, M., Martin, R., Preckel, F., & Goetz, T. (2017). Affect and motivation within and between scool subjects: Development and validation of an integrative structural model of academic self-concept, interest, and anxiety. *Contemporary Educational Psychology, 49*, 46-65.
Goodman, R. (2001). Psychometric properties of the Strenghts and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(11), 1337-1345.
Gut, J., Heckmann, C., Meyer, C.S., Schmid, M., & Grob, A. (2012). Language skills, mathematical thinking, and achievement motivation in children with ADHD, disruptive disorders, and normal controls. *Learning and Individual Differences, 22*(3), 375-379.
Hirschfield, P.J., & Gasper, J. (2011). The relationship between school engagement and delinquency in late childhood and early adolescence. *Journal of Youth and Adolescence, 40*(1), 3-22.
Hoge, R.D., Meginbir, L., Khan, Y., & Weatherall, D. (1985). A multitrait-multimethod analysis of the Preschool Behavior Questionnaire. *Journal of Abnormal Child Psychology, 13*, 119-127.
Jarrett, M., Siddiqui, S., Lochman, J., & Qu, L. (2014). Internalizing problems as a predictor of change in externalizing problems in at-risk youth. *Journal of Clinical Child & Adolescent Psychology, 43*(1), 27-35.
Junod, R.E.V., DuPaul, G.J., Jitendra, A.K., Volpe, R.J., & Cleary, K.S. (2006). Classroom observations of students with and without ADHD. *Journal of School Psychology, 44*(2), 87-104.
Konold, T., Cornell, D., Shukla, K., & Huang, F. (2017). Racial/ethnic differences in perceptions of school climate and its association with student engagement and peer aggression. *Journal of Youth and Adolescence, 46*(6), 1289-1303.
Kurdi, V. & Archambault, I. (2020). Self-perceptions and engagemet in low socio-economic celementary school students: the moderating effects of immigration status and anxiety. *School Mental Health.* Advanced Online Publication.
Li, Y., Bebiroglu, N., Phelps, E., Lerner, J.V., & Lerner, R.M. (2008). Out-of-school time activity participation, school engagement and positive youth development: Findings from the 4-H Study of Positive Youth Development. *Journal of Youth Development, 3*(3), 1-16.
Li, Y., & Lerner, R.M. (2011). Trajectories of school engagement during adolescence: Implications for grades, depression, delinquency, and substance use. *Developmental Psychology, 47*(1), 233-247.
Little, T.D. (2013). *Longitudinal Structural Equation Modeling*. New Nork, NY: The Guildford Press.
Marsh, H.W., Hau, K.T., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares & J.J. McArdle (Eds.), *Contemporary psychometrics* (pp. 275-340). Erlbaum.
Marsh, H.W., Hau, K.T., Wen, Z., Nagengast, B., & Morin, A.J.S. (2013). Moderation. In T. D. Little (Ed.), *Oxford Handbook of Quantitative Methods*. Oxford University Press.
MEES. (2019). *Diplomation et qualification par commission scolaire au secondaire*. Ministère de l'Éducation et de l'Enseignement supérieur du Québec.
Moilanen, K.L., Shaw, D.S., & Maxwell, K.L. (2010). Developmental cascades: Externalizing, internalizing, and academic competence from middle childhood to early adolescence. *Development and Psychopathology, 22*(3), 635-653.

Morin, A.J.S., Arens, A.K., Maiano, C., Ciarrochi, J., Tracey, D., Parker, P.D., & Craven, R.G. (2017). Reciprocal relationships between teacher ratings of internalizing and externalizing behaviors in adolescents with different levels of cognitive abilities. *Journal of Youth & Adolescence, 46*(4), 801-825.

Morin, A.J.S., Boudrias, J.S., Marsh, H.W., Madore, I., & Desrumaux, P. (2016). Further reflections on disentangling shape and level effects in person-centered analyses: An illustration exploring the dimensionality of psychological health. *Structural Equation Modeling, 23*(3), 438-454.

Muthén, L.K., & Muthén, B.O. (2019). *Mplus User's Guide*. Muthén & Muthén.

Nguyen, M. N., Watanabe-Galloway, S., Hill, J. L., Siahpush, M., Tibbits, M. K., & Wichman, C. (2019). Ecological model of school engagement and attention-deficit/hyperactivity disorder in school-aged children. *European Child and Adolescent Psychiatry, 28*(6), 795-805.

Nivard, M.G., Lubke, G.H., Dolan, C.V., Evans, D.M., St Pourcain, B., Munafo, M., & Middeldorp, C.M. (2017). Joint developmental trajectories of internalizing and externalizing disorders between childhood and adolescence. *Development and Psychopathology, 29*(3), 919-928.

OECD. (2018). *Equity in Education: Breaking Down Barriers to Social Mobility*. Paris: PISA, OECD.

Olivier, E., & Archambault, I. (2017). Hyperactivity, inattention, and student engagement: The protective role of relationships with teachers and peers. *Learning and Individual Differences, 59*, 86-95.

Olivier, E., Archambault, I., & Dupéré, V. (2018). Boys' and girls' latent profiles of behavior and social adjustment in school: Longitudinal links with later student behavioral engagement and academic achievement? *Journal of School Psychology, 69*, 28-44.

Olivier, E., Archambault, I., & Dupéré, V. (2020). Do needs for competence and relatedness mediate the risk of low engagement of students with behavior and social problem profiles? *Learning & Individual Differences, 78*.

Papachristou, E., & Flouri, E. (2019). The codevelopment of internalizing symptoms, externalizing symptoms, and cognitive ability across childhood and adolescence. *Development and Psychopathology*, 1-15.

Plamondon, A., & Martinussen, R. (2019). Inattention symptoms are associated with academic achievement mostly through variance shared with intrinsic motivation and behavioral engagement. *Journal of Attention Disorders, 23*(14), 1816-1828.

Radloff, L.S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385-401.

Randazzo, K.V., Landsverk, J., & Ganger, W. (2003). Three informants' reports of child behavior: Parents, teachers, and foster parents. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(11), 1343-1350.

Reitz, E., Deković, M., & Meijer, A.M. (2005). The structure and stability of externalizing and internalizing problem behavior during early adolescence. *Journal of Youth and Adolescence, 34*(6), 577-588.

Searle, A. K., Miller-Lewis, L. R., Sawyer, M. G., & Baghurst, P. A. (2013). Predictors of children's kindergarten classroom engagement: Preschool adult–child relationships, self-concept, and hyperactivity/inattention. *Early Education & Development, 24*(8), 1112-1136.

Skinner, E.A., Kindermann, T.A., Connell, J.P., & Wellborn, J.G. (2009). Engagement and disaffection as organizational construct in the dynamics of motivational development. In K.R. Wentzel & A. Wigfield (Eds.), *Handbook of Motivation at School* (pp. 223-246). Routledge.

Smith, S.R. (2007). Making sense of multiple informants in child and adolescent psychopathology: A guide for clinicians. *Journal of Psychoeducational Assessment*, 25(2), 139–149.

Tardif-Grenier, K., Olivier, E., & Archambaul, I. (2019). Is there an immigrant paradox in Canadian elementary students' behavioral and social adjustment? *International Journal of School & Educational Psychology*, Advanced online publication.

Tian, Y., Bian, Y., Han, P., Gao, F., & Wang, P. (2017). Class collective efficacy and class size as moderators of the relationship between junior middle school students' externalizing behavior and academic engagement: A multilevel study. *Frontiers in Psychology, 8*, 1-9.

Trembaly, R.E., Desmarais-Gervais, L., Gagnon, C., & Charlebois, P. (1987). The Preschool Behavior Questionnaire: Stability of its factor structure between cultures, sexes, ages and socioecconomic classes. *Internationnal Journal of Behavioral Development, 10*, 467-484.

UNESCO. (2020). *World Inequality Database on Education*. Montréal, QC: UNESCO Institute for Statistics.

Vasey, M., Bosmans, G., & Ollendick, T. (2014). The developmental psychopathology of anxiety. In M. Lewis & K. Rudolph (Eds.), *Handbook of Developmental Psychopathology, 3rd Ed.* (pp. 543-560). New York,: Springer.

Wang, F.L., Eisenberg, N., Valiente, C., & Spinrad, T.L. (2016). Role of temperament in early adolescent pure and co-occurring internalizing and externalizing problems using a bifactor model. *Development and Psychopathology, 28*(4pt2), 1487-1504.

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

Wang, M.T., & Eccles, J.S. (2012). Adolescent behavioral, emotional, and cognitive engagement trajectories in

school and their differential relations to educational success. *Journal of Research on Adolescence, 22*(1), 31-39.
Wang, M.T., Fredricks, J., Ye, F., Hofkens, T., & Linn, J.S. (2019). Conceptualization and assessment of adolescents' engagement and disengagement in school. *European Journal of Psychological Assessment, 35*(4), 592-606.
Wang, M.T., & Fredricks, J. (2014). The reciprocal links between school engagement, youth problem behaviors, and school dropout during adolescence. *Child Developement, 85*(2), 722-737.
Wang, M.T., Fredricks, J.A., Ye, F., Hofkens, T.L., & Linn, J.S. (2016). The Math and Science Engagement Scales: Scale development, validation, and psychometric properties. *Learning and Instruction, 43*, 16-26
Wang, M. T., & Peck, S. C. (2013). Adolescent educational success and mental health vary across school engagement profiles. *Developmental Psychology, 49*(7), 1266-1276.
Wigfield, A. & Eccles, J S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81.
Willner, C.J., Gatzke-Kopp, L.M., & Bray, B.C. (2016). The dynamics of internalizing and externalizing comorbidity across the early school years. *Development and Psychopathology, 28*(4pt1), 1033-1052.

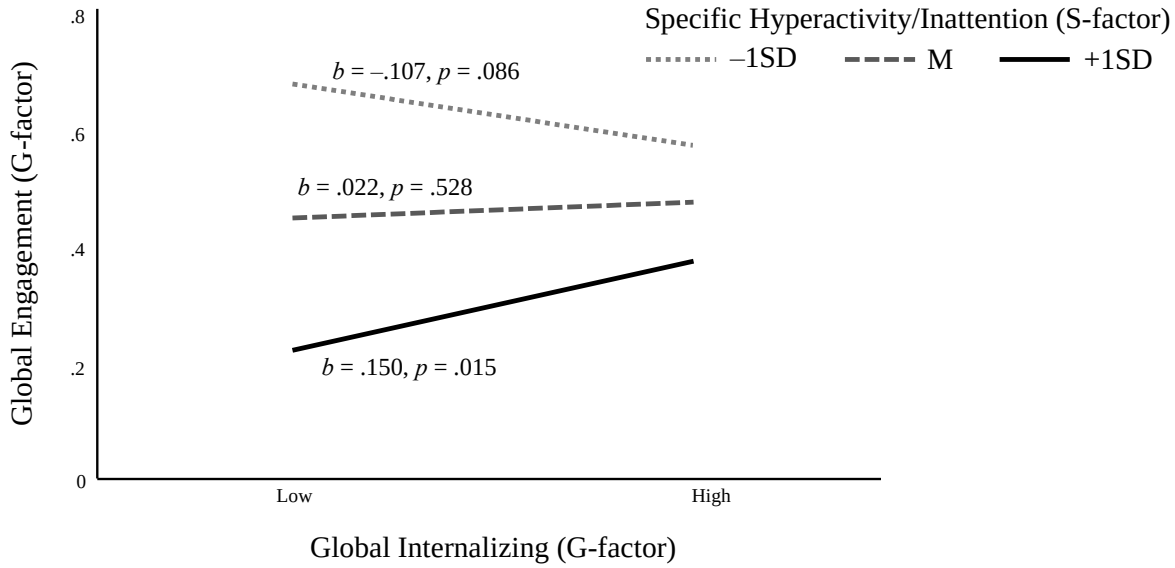

Figure 1. Results from the global internalizing behaviors (G-factor) × specific hyperactivity/inattention interaction (S-factor) in the prediction of global engagement levels (G-factor) in the secondary school sample.

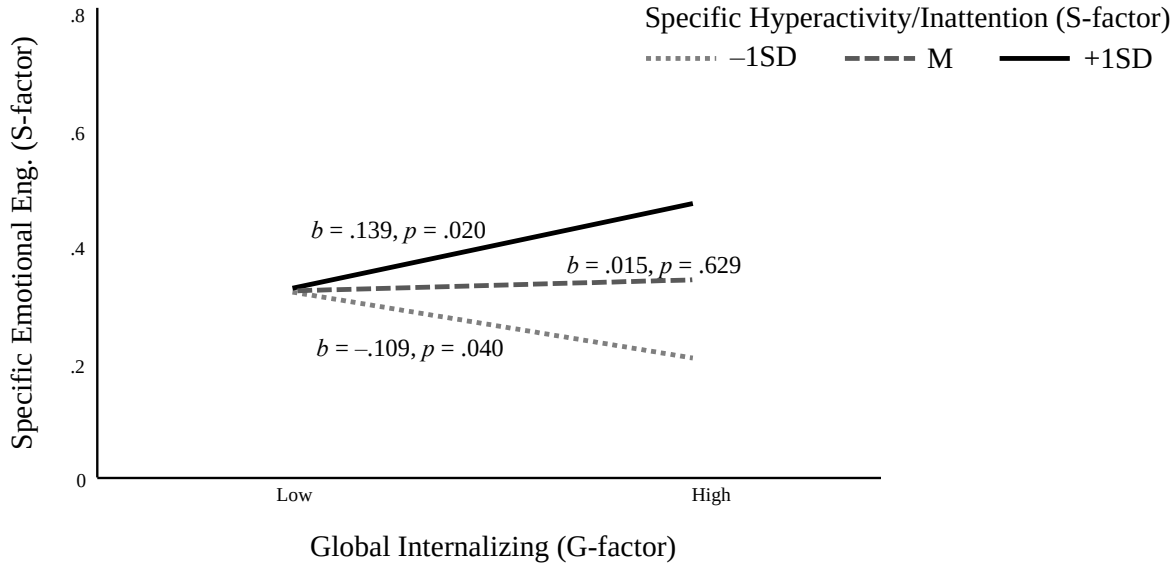

Figure 2. Results from the global internalizing behaviors (G-factor) × specific hyperactivity/inattention interaction (S-factor) in the prediction of specific emotional engagement levels (S-factor) in the secondary school sample.

Table 1

Literature Review

	Age /Dev. Period	Behavior problems	Dimensions of engagement or related measures	Results[1]	Comparisons[2]	
					Sex	Age /Dev.
Externalizing behaviors						
Childhood						
Archambault et al. (2017)	Grade 3 and 4	O/D (tr)	BE and EE (sr; lang)	O/D → lower BE, EE	Yes: no differences	No.
Demaray & Jenkins (2011)	Grade 3 to 5	H/I (pr)	Eng. (academic enabler) (tr; nss)	H/I group → lower E	Yes: no differences	No.
Junod et al. (2006)	Grade 1 to 4	H/I (diagnosis)	BE (om; nss)	H/I group → lower BE	No.	No.
Moilanen et al. (2010)	6 to 12 y.o.	O/D (pr)	BE (tr; nss)	O/D → lower BE	N/A (boys only sample)	No.
Olivier & Archambault (2017)	Grade 4 to 6	H/I (tr)	BE, EE, and CE (sr; lang)	I and H→ lower BE	Yes: no differences	No.
Plamondon & Martinussen (2019)	Grade 1 to 4	Inattention (pr, tr)	BE (tr; nss)	I → lower BE	No.	No.
Volpe et al. (2006)	Grade 1 to 4	H/I (diagnosis)	Eng. (academic enabler) (tr; nss)	H/I group → lower BE	No.	No.
Adolescence						
Tian et al. (2017)	Grade 7 to 9	O/D (sr)	Eng. (vigor, absorption, and dedication) (sr; nss)	O/D → lower Eng.	No.	No.
Wang & Fredricks (2014)	12-17 y.o.	O/D (sr)	Global Eng. (BE, EE, and CE) (sr; nss)	O/D → lower Global. Eng.	No.	No.
Childhood and adolescence						
Nguyen et al. (2019)	6-17 y.o.	H/I (diagnosis)	Eng. (two items: one BE and one EE) (pr; nss)	H/I group → lower Eng.	No.	No.
Internalizing behaviors						
Childhood						
Kurdi & Archambault (2020)	Grade 4 to 6	Anx. (sr)	BE, EE, and CE (sr; math)	No direct association.	Yes: no differences	No.
Adolescence						
Derdikman-Eiron et al. (2011)	13-19 y.o.	Int. (sr)	BE (sr; nss)	Int. → lower BE	Yes: larger effect for boys	No.
Dorio et al. (2019)	Grade 6 to 8	Dep. (sr)	Global Eng. (BE, EE, and CE) (sr; nss)	Dep. → lower Glo. E	Yes: larger effect for girls	No.
Fiorilli et al. (2017)	14-16 y.o.	Dep. (sr)	Eng. (vigor, absorption, and dedication) (sr; nss)	Dep. → lower BE	No.	No.
Garvik et al. (2014)	15-18 y.o.	Dep. (sr)	BE (sr; nss)	Dep. → lower BE	Yes: no differences	No.
Wang & Peck (2013)	Grade 9 and 11	Dep (sr)	BE, EE, and CE (sr; nss)	Dep. → low EE profile Dep. → low BE+EE+CE profile	No.	No.
Wang et al. (2015)	Grade 9 to 11	Dep. (sr)	EE (sr; nss)	Dep. → lower EE trajectory	No.	No.
Externalizing and internalizing behaviors						
Childhood						
Baker et al. (2008)	Kinder. to grade 5	Ext. and Int. (tr)	BE (tr; nss)	Int. → lower BE	No.	No.
Olivier et al. (2018; 2020)	Grade 5 and 6	Ext. (tr) and Int. (sr)	BE (tr and sr), EE (sr), and CE (sr) (lang; math)	Girls: (no Int. + Ext. profile) Ext. profile → lower BE, EE, CE Int. profile → lower BE, EE, CE Boys: (no Int. profile) Ext. profile: not associated Int. + Ext. profile → lower BE, EE	Yes. Diff. listed in results.	No.
Searle et al. (2013)	Kinder. and Grade 1	H/I, O/D, Dep. (pr and tr)	Global Eng. (BE, EE, and CE) (tr; ; nss)	H/I → lower Glo. E	No.	No.
Adolescence						
Curhan et al. (2020)	Grade 7 (11-13 y.o.)	Ext. and Int. (sr, pr, and tr)	EE (sr; nss)	Int → lower EE	No.	No.
Li et al. (2008)	Grade 8	O/D and Dep. (sr)	BE and EE (sr; nss)	O/D → lower EE Dep. → lower BE, EE	No.	No.
Li & Lerner (2011)	Grades 5 to 8	O/D and Dep. (sr)	BE and EE (sr; nss)	O/D and Dep. → lower BE, EE	No.	No.

Note. H/I = hyperactivity/inattention; O/D = opposition and defiance; Ext. = externalizing behaviors comprising H/I and O/D; Anx. = anxiety; Dep. = Depressive symptoms; Int. = internalizing behaviors comprising anx. and dep.; Eng. = engagement; BE = behavioral engagement; EE = emotional engagement; CE = cognitive engagement; sr = student-rated; tr = teacher-rated; pr = parent-rated; om = observational measure; math = mathematics class engagement; lang = language class engagement; nss = non-subject specific engagement.

[1]: Only significant results are reported.

[2]: These columns list the comparisons of associations between behavior problems and engagement, but not of levels of behavior problems and engagement.

Table 2

Correlations between Factor Scores and Covariates

		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.
1.	Sex		.13**	–.02	–	.12**	–.02	.01	–.07*	–.18**	.29**	.06	.09**	.08*	.15**	–.07*	.12**
2.	Grade level	–.01		.00	–	–.32**	.11**	.08*	–.01	.04	.02	–.03	.05	–.14**	–.21**	–.09**	.02
3.	Family (0=intact)	–.01	.05		–	–.06*	–.16**	.02	.08*	–.04	.04	.01	.10**	–.08*	–.04	–.07*	–.01
4.	Achievement	.10**	–.17**	–.10**		–	–	–	–	–	–	–	–	–	–	–	–
5.	Retention (0=yes)	–	–	–	–		.10**	.01	.03	–.01	.03	.00	–.09**	.07*	.13**	.00	.05
6.	Immigrant (0=no)	–	–	–	–	–		.22**	–.07*	.02	–.05	–.18**	.03	–.02	–.13**	–.06	.05
7.	G-Ext	–.22**	.03	.05	–.35**	–	–		.28**	.11**	.17**	–.07*	.20**	–.16**	–.18**	–.08*	–.04
8.	S-H/I	.03	.03	–.01	–.28**	–	–	.11**		.04	.11**	.03	.10**	–.21**	–.17**	–.01	–.08*
9.	S-O/D.	–.02	.04	.01	–.00	–	–	.13**	–.10**		–.08*	–.06	.02	–.08*	–.12**	.01	–.06
10.	G-Int	.10*	.01	.12**	–.10**	–	–	.11**	.12**	.04		–.03	.06	.01	–.01	–.03	.00
11.	S-Anx	.08*	.09*	.01	.03	–	–	–.01	.01	–.03	.27**		–.03	.05	.09**	–.01	.02
12.	S-Dep	–.05	–.01	.08	–.03	–	–	.09*	.07	.07	.29**	–.31**		–.19**	–.08**	–.15**	.00
13.	G-Eng	.27**	–.18**	–.08*	.19**	–	–	–.32**	–.05	–.04	–.20**	–.10*	–.11**		.10**	.33**	.44**
14.	S-BE	.05	.00	–.01	.12**	–	–	–.21**	–.04	.02	–.08*	–.06	–.01	.21**		–.13**	–.17**
15.	S-EE	.21**	–.19**	–.01	.18**	–	–	–.08*	–.04	–.10	–.01	–.08*	–.02	.29**	–.21**		–.37**
16.	S-CE	.02	–.04	–.01	–.02	–	–	.00	.03	–.06	–.13**	.04	–.09*	.28**	–.32**	–.24**	

Note. Correlations in the elementary school sample are displayed below the diagonal, and those in the secondary school sample are displayed above the diagonal. Bifactor correlations are not exactly zero as these involve factor scores; G: Global factor from a bifactor solution; S: Specific factor from a bifactor solution; Ext: Externalizing behaviors; H/I: Hyperactivity/Inattention; O/D: Oppositional/Defiant; Int: Internalizing behaviors; Anx: Anxiety; Dep: Depressive symptoms; Eng: Engagement; BE: Behavioral engagement; EE: Emotional engagement; CE: Cognitive engagement.

* $p \leq .05$; ** $p \leq .01$.

Table 3

Results from the Predictive (Structural) Models

	Global Eng.			Specific Behavioral Eng.			Specific Emotional Eng.			Specific Cognitive Eng.		
	b	s.e.	β	b	s.e.	β	b	s.e.	β	b	s.e.	β
Elementary school sample												
Covariates Model												
Sex (0=boy)	.426	.060**	.257	.051	.044	.044	.262	.048**	.202	.024	.047	.019
Grade level	–.093	.028**	–.129	.004	.019	.008	–.093	.022**	–.164	–.008	.021	–.014
Family (0=intact)	–.100	.064	–.057	.002	.047	.002	.022	.052	.016	–.008	.050	–.006
Achievement	.120	.030**	.153	.066	.021**	.122	.090	.023**	.147	–.013	.024	–.022
Full Model												
Sex (0=boy)	.363	.054**	.220				.298	.045**	.229			
Grade level	–.091	.025**	–.126				–.090	.019**	–.159			
Family (0=intact)												
Achievement	.052	.031*	.067	.021	.021	.039	.090	.025**	.146			
G-Ext	–.256	.046**	–.232	–.155	.035**	–.203	.024	.036	.028	.018	.032	.022
S-H/I	.028	.067	.017	.060	.052	.053	–.012	.055	–.009	–.059	.052	–.048
S-O/D	.023	.061	.016	.009	.048	.009	–.014	.047	.012	.037	.047	.033
G-Int	–.124	.038**	–.143	–.022	.028	–.037	.015	.034	.023	–.089	.030**	–.138
S-Anx	–.151	.067*	–.095	–.056	.052	–.051	–.139	.056**	–.111	.074	.054	.063
S-Dep	–.107	.062	–.070	–.001	.052	.001	–.053	.057	–.044	–.037	.057	–.032
Secondary school sample												
Covariates Model												
Sex (0=boy)	–.114	.053*	–.067	–.177	.042**	–.127	.116	.048*	.076	–.172	.044**	–.123
Grade level	–.212	.052**	–.126	–.194	.042**	–.140	–.178	.047**	–.116	.066	.043	.047
Family (0=intact)	–.142	.055**	–.077	–.089	.048	–.059	–.145	.053**	–.086	.010	.049	.007
Retention (0=no)	.142	.110	.046	.252	.088**	.100	–.007	.087	–.003	.097	.082	.038
Immigrant (0=no)	–.062	.054	–.036	–.189	.043**	–.134	–.124	.049*	–.079	.087	.044*	.061
Full Model												
Sex (0=boy)	–.094	.054	.055	–.152	.046**	.109	.101	.052*	–.056	–.171	.046**	.121
Grade level	–.266	.045**	–.158	–.196	.041**	–.142	–.164	.044**	–.107			
Family (0=intact)	–.092	.049	–.050				–.101	.046**	–.083			
Retention (0=no)				.275	.076**	.101						
Immigrant (0=no)				–.143	.042**	–.108	–.101	.048*	.064	.097	.042*	.068
G-Ext	–.129	.043**	–.107	–.105	.036**	–.106	–.059	.038	–.054	–.035	.035	–.035
S-H/I	–.221	.043**	–.165	–.149	.036**	–.136	.015	.039	.013	–.057	.035	–.051
S-O/D	–.052	.065	–.026	–.124	.051**	–.075	.021	.063	.012	–.040	.056	–.024
G-Int	.041	.034	.043	–.013	.028	.016	.017	.031	.019	–.019	.027	.024
S-Anx	.045	.038	.035	.056	.031	.052	–.041	.038	–.034	.024	.037	.022
S-Dep	–.148	.033**	–.144	–.033	.027	–.039	–.115	.031**	–.123	.006	.028	.007

Note. b: Unstandardized regression coefficient; s.e. Standard error of the coefficient; β: Standardized regression coefficient; G: Global factor from a bifactor solution; S: Specific factor from a bifactor solution; Ext: Externalizing behaviors; H/I: Hyperactivity/Inattention; O/D: Oppositional/Defiant; Int: Internalizing behaviors; Anx: Anxiety; Dep: Depressive symptoms; Eng: Engagement.

$^{*}p < .05$; $^{**}p < .01$.

Appendix 1
Measurement models in the elementary school sample

Analyses

We conducted CFA and bifactor-CFA using the robust weight least square (WLSMV) estimator, which outperforms Maximum Likelihood estimation with ordinal rated using five or fewer response categories and/or asymmetric response thresholds (Finney & DiStefano, 2013). In these models, there were no missing data on student-reported measures as per the specifications of the computerized questionnaire. There were 3.58% of missing data on teacher-reported measures, which were handled using the missing data procedures implemented in Mplus for WLSLMV estimation (Asparouhov & Muthén, 2010), thus allowing the models to be estimated using all of the available information without relying either on deletion or imputation procedures.

To select the optimal model (CFA or bifactor-CFA), we relied on Morin et al. (2020) recommendations stating that a bifactor-CFA solution should be favored over a CFA solution when it results in "(a) an improved level of fit the data; (b) a well-defined G-factor; (c) at least some reasonably well-defined S-factors" (p. 14). We also assessed the composite reliability of each factor using omega (ω; McDonald, 1970), which takes into account the strength of the associations between items and all constructs, as well as item-specific measurement error, and has been recommended to be appropriate for CFA and bifactor-CFA measurement (Morin et al., 2020). After selecting the best model, we performed tests of measurement invariance across sexes following the steps described by Meredith (1993), as adapted to categorical indicators (Morin et al., 2016; Morin et al., 2011): (i) configural invariance (same model with no additional constraint); (ii) equal factor loadings (weak invariance); (iii) equal response thresholds (strong invariance); (iv) equal item uniquenesses (strict invariance); (v) equal factor variances-covariances (CFA) or variances (bifactor-CFA, due to the orthogonality of the model); (vi) equal latent means. The last two steps are not required to establish measurement invariance, but were investigated for descriptive purposes.

Results

Externalizing Behaviors. For externalizing behaviors, the bifactor-CFA model (χ^2 = 109.239, df = 15, p < .01; RMSEA = .008; CFI = .990; TLI = .975) resulted in an excellent level of fit to the data that proved to be substantially greater than that of the CFA solution (χ^2 = 275.896, df = 23, p < .01; RMSEA = .106; CFI = .972; TLI = .956). The parameter estimates from this model are reported in Table S1, and revealed that the G- and S-factors were generally well-defined, although the Hyperactivity/Inattention S-factor was more weakly defined (|λ| = .014-.523; ω= .688) than the Externalizing behaviors G-factor (|λ| = .612-.929; ω= .950) and the Oppositional/Defiant S-factor (|λ| = .402-.508; ω= .740). Importantly, given the high correlation between the Hyperactivity/Inattention and Oppositional/Defiant factors in the CFA solution (r = .787, p < .01), the orthogonal bifactor-CFA solution also appeared to solve the issue of conceptual redundancy between factors, and was retained as our final solution. Results from the tests of measurement invariance, reported in Table S2, supported the invariance of the factor loadings, response thresholds, item uniquenesses, and factor variances across sexes. However, these results also revealed latent mean difference between groups, suggesting that, whereas boys and girls displayed similar means on the Hyperactivity/Inattention and Oppositional/Defiant S-factors, boys' latent mean on the Externalizing behaviors G-factor were .628 SD higher than girls' latent mean.

Internalizing Behaviors. For internalizing behaviors, the results showed that the bifactor-CFA model (χ^2 = 35.123, df = 9, p < .01; RMSEA = .064; CFI = .991; TLI = .980) resulted in an excellent level of fit to the data that proved to be substantially greater than that of the CFA solution (χ^2 = 84.335, df = 13, p < .01; RMSEA = .088; CFI = .977; TLI = .962). The parameter estimates from this model are reported in Table S1, and revealed that the G- and S- factors were generally well-defined, although the Anxiety S-factor was more weakly defined (|λ| = .127-.526; ω= .555) than the Internalizing behaviors G-factor (|λ| = .646-.800; ω= .918) and the Depression S-factor (|λ| = .414-.531; ω= .703). As for externalizing behaviors, the bifactor-CFA solution solved the conceptual redundancy suggested by the high correlation observed between the Anxiety and Depression factors in the CFA solution (r = .786, p < .01), and was retained as our final solution. Results from the tests of measurement invariance, reported in Table S2, supported the invariance of the factor loadings, response thresholds, item uniquenesses, factor variances, and latent means across sexes. Regarding the more weakly defined Anxiety S-factor, it is important to keep in mind that, due to the fact that bifactor model divide true score (i.e., reliable) variance across two factors, it often happens that some indicators are found to retain little specific variance once the variance explained by the G-factor is taken into account. More precisely, the present results suggest that ratings of anxiety seem to retain a limited amount of specificity once the true score variance explained by global levels of internalizing behaviors are extracted from them. This observation is consistent with the fact anxiety and depression share a high level of comorbidity and that, when they co-occur, anxiety is more often a precursor of depression than depression a precursor of anxiety (e.g., Barrocas & Hankin, 2011; Rice & Thapar, 2009). However, with an ω= .555, even this more weakly defined S-factors can be considered to retain enough specificity to be meaningful (Morin et al., 2020).

Student Engagement. For engagement, the results showed that the bifactor-CFA model (χ^2 = 55.873, df = 18, p < .01; RMSEA = .055; CFI = .994; TLI = .988) resulted in an excellent level of fit to the data that proved to be substantially greater than that of the CFA solution (χ^2 = 130.995, df = 24, p < .01; RMSEA = .080; CFI = .983; TLI = .975). The parameter estimates from this model are reported in Table S1, and revealed that the G- and S- factors were generally well-defined, although the Behavioral engagement S-factor was more weakly defined (|λ| = .122-.506; ω= .576) than the engagement G-factor (|λ| = .531-.792; ω= .915), the Emotional engagement S-factor (|λ| = .452-.504; ω= .662), and the Cognitive engagement S-factor (|λ| = .310-.549; ω= .569). For the Behavioral engagement S-factor, the fact that it was mainly defined by items referring to listening to the teachers and following instructions, rather than effort, suggest that this factor represents student behavioral compliance in class. As for externalizing and internalizing profiles, the bifactor-CFA solution appeared to solve the conceptual redundancy suggested by the high correlation observed between the Behavioral, Emotional, and Cognitive engagement factors in the CFA solution (r = .652 to .716, p < .01). This bifactor-CFA model was thus retained as our final solution. Results from the tests of measurement invariance, reported in Table S2, supported the invariance of the factor loadings, response thresholds, item uniquenesses, and factor variances across sexes. The results also revealed latent mean differences, suggesting that, whereas boys and girls had similar means on the Behavioral, Emotional and Cognitive engagement S-factors, girls had higher latent means on the Engagement G-factor (+.639 SD) relative to boys.

References

Asparouhov, T., & Muthén, B.O. (2010). *Weighted least square estimation with missing data*: www.statmodel.com/download/GstrucMissingRevision.pdf.

Finney, S.J., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G.R. Hancock & R.O. Mueller (Eds.), *Quantitative methods in education and the behavioral sciences: Issues, research, and teaching. Structural equation modeling: A second course* (pp. 439-492). Charlotte, NC: IAP Information Age Publishing.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525-543.

McDonald, R.P. (1970). The theoretical foundations of principal factor analysis, canonical factor analysis, and alpha facctor analysis. *British Journal of Mathematical and Statistical Psychology, 23*(1), 1-21.

Morin, A.J.S., Arens, A.K., Tran, A., & Caci, H. (2016). Exploring sources of construct-relevant multidimensionality in psychiatric measurement: A tutorial and illustration using the Composite Scale of Morningness. *International Journal of Methods in Psychiatric Research, 25*(4), 277-288.

Morin, A.J.S., Myers, N.D., & Lee, S. (2020). Modern factor analytic techniques: Bifactor models, exploratory structural equation modeling (ESEM) and bifactor-ESEM. In G. Tenenbaum & R.C. Eklund (Eds.), *Handbook of Sport Psychollogy, 4th Edition*. London, UK: Willey.

Morin, A.J.S., Moullec, G., Maïano, C., Layet, L., Just, J.L., & Ninot, G. (2011). Psychometric properties of the Center for Epidemiologic Studies Depression Scale (CES-D) in French clinical and nonclinical adults. *Epidemiology and Public Health, 59*(5), 327-340.

Table S1

Standardized Factor Loadings (λ) and Uniquenesses (δ) from the Bifactor Confirmatory Factor Analytic Results from the Elementary School Sample.

Items	G-Ext λ	S-H/I λ	S-O/D λ	G-Int λ	S-Anx λ	S-Dep λ	G-Eng λ	S-BE λ	S-EE λ	S-CE λ	δ
Externalizing behaviors											
Overactive	.929	–.287									.055
Fidgeting	.912	–.320									.067
Distracted	.779	.420									.216
Thinks ahead (R)	.692	*–.014*									.504
Finishes tasks (R)	.612	.523									.266
Loses temper	.697		.518								.255
Well behaved (R)	.666		.487								.182
Fights	.704		.445								.307
Lies, cheats	.613		.402								.463
Internalizing behaviors											
Fearful				.669	.526						.276
Nervous				.746	.245						.384
Worried about others				.800	*–.127*						.344
Worried about school				.704	.399						.345
Sad				.721		.490					.240
Unhappy				.708		.531					.217
Does not have fun				.646		.414					.412
Engagement											
Follow instructions							.774	.506			.144
Listen							.792	.485			.138
Make efforts							.597	*–.122*			.629
Interested							.698		.504		.259
Appreciative							.675		.452		.340
Bored (R)							.531		.494		.474
Check for errors							.603			.549	.335
Tries to understand							.691			.436	.333
Uses strategies							.547			.310	.604
Composite reliability (ω)	.950	.688	.740	.918	.555	.703	.915	.576	.662	.569	

Note. G: Global factor from a bifactor solution; S: Specific factor from a bifactor solution; Ext: Externalizing behaviors; H/I: Hyperactivity/Inattention; O/D: Oppositional/Defiant; Int: Internalizing behaviors; Anx: Anxiety; Dep: Depression; Eng: Engagement; BE: Behavioral engagement; EE: Emotional engagement; CE: Cognitive engagement; ω: omega coefficient of composite reliability (McDonald, 1970).

Non-statistically significant results ($p \leq .05$) are marked in italics.

Table S2

Measurement and structural invariance tests across sex in the elementary school sample.

Model	χ^2	*df*	CFI	TLI	RMSEA	RMSEA 90% CI	$\Delta\chi^2$	Δdf	ΔCFI	ΔTLI	ΔRMSEA
					Measurement models						
Bifactor-CFA Externalizing behaviors by sex											
1. Configural invariance	118.257*	31	.986	.967	.090	.073–.107	–	–	–	–	–
2. Weak invariance	177.490*	50	.980	.971	.085	.072–.099	74.889*	19	–.006	.004	–.005
3. Strong invariance	174.923*	55	.981	.975	.079	.066–.092	1.949	5	.001	.004	–.006
4. Strict invariance	183.193*	64	.981	.978	.073	.061–.085	18.163*	9	.000	.003	–.006
5. Latent variance invariance	138.602*	68	.989	.988	.054	.041–.067	5.596	4	.008	.010	–.019
6. Latent mean invariance	259.599*	72	.970	.970	.086	.075–.098	80.839*	4	–.019	–.018	.032
6. Latent mean invariance–partial	145.248*	71	.988	.988	.055	.042–.067	8.300*	3	–.001	.000	.001
Bifactor-CFA Internalizing behaviors by sex											
1. Configural invariance	33.571*	14	.994	.982	.063	.036–.091	–	–	–	–	–
2. Weak invariance	36.944*	25	.996	.993	.040	.010–.063	10.273	11	.002	.011	–.023
3. Strong invariance	46.300*	29	.995	.992	.041	.016–.036	7.538	4	–.001	–.001	.001
4. Strict invariance	55.026*	36	.994	.993	.039	.015–.058	10.834	7	–.001	.001	–.002
5. Latent variance invariance	49.640	39	.997	.996	.028	.000–.049	3.146	3	.003	.003	–.011
6. Latent mean invariance	71.520*	42	.991	.991	.045	.026–.062	14.610*	3	–.006	–.005	.017
Bifactor–CFA Engagement by sex											
1. Configural invariance	92.051*	36	.990	.981	.067	.050–.084	–	–	–	–	–
2. Weak invariance	103.706*	50	.991	.987	.055	.040–.070	18.832	14	.001	.006	–.012
3. Strong invariance	127.806*	73	.991	.991	.046	.033–.059	31.641	23	.000	.004	–.009
4. Strict invariance	138.339*	82	.990	.992	.044	.031–.057	16.553	9	–.001	.001	–.002
5. Latent variance invariance	117.347*	86	.995	.996	.032	.015–.046	3.479	4	.005	.004	–.012
6. Latent mean invariance	374.194*	90	.952	.961	.095	.085–.105	100.768*	4	–.043	–.035	.063
6. Latent mean invariance–partial	146.591*	89	.990	.992	.043	.030–.055	22.820*	3	–.005	–.004	.011
Predictive (Structural) Model Invariance by sex											
1. Configural invariance	3.084	6	1.000	1.057	.000	.000-.043	–	–	–	–	–
2. Regression invariance	40.315	41	1.000	1.002	.000	.000-.034	37.114	35	.000	–.055	.000
3. Intercept invariance	97.100*	45	.919	.863	.055	.040-.070	57.252*	4	–.081	–.139	.055
3. Intercept invariance–partial	41.681	43	1.000	1.004	.000	.000-.033	1.323	2	.000	.002	.000
4. Residuals invariance	53.023	47	.991	.985	.018	.000-.040	10.662*	4	–.009	–.019	.018
4. Residuals invariance–partial	47.938	46	.997	.995	.011	.000-.036	6.147	3	–.003	–.009	.011

Note. χ^2: Chi square test of model fit and associated degrees of freedom (*df*); CFI: Comparative Fit Index; TLI: Tucker-Lewis Index; RMSEA: Root Mean Square Error of Approximation and 90% Confidence Interval (CI); Δ: Change according to the previous retained model; $\Delta\chi^2$: Chi square difference test calculated with the Mplus DIFFTEST option for the measurement invariance model and with the Satorra-Bentler correction for the predictive (structural) models.

*$p < .05$.

Appendix 2
Measurement models in the secondary school sample

Analyses

We followed the same procedure as in Study 1. Given that Study 2 was conducted among a multiethnic sample, we also tested for the measurement invariance and predictive equivalence as a between groups of immigrant (1st and 2nd generation) and non-immigrant students (3rd generation+).

Results

Externalizing Behaviors. For externalizing behaviors, the results showed that the bifactor-CFA model (χ^2 = 34.147, df = 22, $p < .05$; RMSEA = .023; CFI = .998; TLI = .995) resulted in an excellent level of fit to the data that proved to be substantially greater than that of the CFA solution (χ^2 = 175.133, df = 31, $p < .01$; RMSEA = .067; CFI = .971; TLI = .958). The parameter estimates from this model are reported in Table S3, and revealed that the G- and S- factors were generally well-defined, although the Oppositional/Defiant S-factor was more weakly defined (|λ| = .045-.439; ω= .563) than the Externalizing behaviors G-factor (|λ| = .353-.915; ω= .915) and the Hyperactivity/Inattention S-factor (|λ| = .092-.722; ω= .742). As in Study 1, this model solved the issue of conceptual redundancy suggested by the observation of a high correlation between the Hyperactivity/Inattention and Oppositional/Defiant factors in the CFA solution (r = .771, $p < .01$), and was retained as our final solution. Results from the tests of measurement invariance, reported in Table S4, supported the invariance of the factor loadings, response thresholds, item uniquenesses, and factor variances between sexes and immigration status. These results also supported the invariance of the latent means across sexes, but revealed mean differences between immigration status. Thus, whereas immigration status had no impact on the Hyperactivity/Inattention and Oppositional/Defiant S-factors, the mean on the Externalizing behaviors G-factor was .450 SD higher among immigrant relative to non-immigrants.

Internalizing Behaviors. For internalizing behaviors, the results showed that the bifactor-CFA model (χ^2 = 309.318, df = 75, $p < .01$; RMSEA = .056; CFI = .990; TLI = .986) resulted in an excellent level of fit to the data that proved to be substantially greater than that of the CFA solution (χ^2 = 519.014, df = 89, $p < .01$; RMSEA = .069; CFI = .982; TLI = .979). The parameter estimates from this model are reported in Table S3, and revealed that the G- and S- factors were generally well-defined, although the Anxiety S-factor was more weakly defined (|λ| = .063-.447; ω= .546) than the Internalizing behaviors G-factor (|λ| = .544-.824; ω= .952) and the Depression S-factor (|λ| = .625-.740; ω= .929), thus replicating the results from Study 1 (see our earlier discussion of these results for additional details). As in Study 1, this model solved the issue of conceptual redundancy suggested by the observation of a high correlation between the Anxiety and Depression factors in the CFA solution (r = .630, $p < .01$), and was retained as our final solution. Results from the tests of measurement invariance, reported in Table S4, supported the invariance of the factor loadings, response thresholds, item uniquenesses, and factor variances across sexes and immigration status. These results also supported the invariance of the latent means across immigration status, but revealed mean differences between sexes. Indeed, whereas sex had no impact on the Anxiety and Depression S-factors, girls' mean was higher on the Internalizing behavior G-factor by .739 SD relative to boys'.

Student Engagement. For engagement, the results showed that the bifactor-CFA model χ^2 = 518.041, df = 102, $p < .01$; RMSEA = .063; CFI = .975; TLI = .967) resulted in an excellent fit to the data that proved to be similar to that of the CFA solution (χ^2 = 523.615, df = 116, $p < .01$; RMSEA = .058; CFI = .976; TLI = .971). The parameter estimates form this model are reported in Table S3, and revealed that the Engagement G-factor (|λ| = .295-.680; ω= .926), as well as the Behavioral Engagement (|λ| = .465-.743; ω= .756), Emotional engagement S-factor (|λ| = .310-.639; ω= .777) and Cognitive engagement (|λ| = .394-.536; ω= .807) S-factors were all well-defined. As in Study 1, this model solved the issue of conceptual redundancy suggested by the observation of the moderate to high correlations observed between the Behavioral, Emotional, and Cognitive engagement factors in the CFA solution (r = .401 to .598, $p < .01$). The bifactor-CFA was thus retained as our final solution, as decision that was also predicated on the need to achieve comparability across studies. Results from the tests of measurement invariance, reported in Table S4, supported the invariance of the factor loadings, response thresholds, item uniquenesses, factor variances, and latent means across sexes and immigration status.

Table S3

Standardized Factor Loadings (λ) and Uniquenesses (δ) from the Bifactor Confirmatory Factor Analytic Results from the Secondary School Sample.

Items	G-Ext λ	S-H/I λ	S-O/D λ	G-Int λ	S-Anx λ	S-Dep λ	G-Eng λ	S-BE λ	S-EE λ	S-CE λ	δ
Externalizing behaviors											
Overactive	.561	.722									.165
Distracted	.621	.475									.389
Thinks ahead (R)	.353	*–.092*									.253
Fidgeting	.551	.597									.340
Finishes tasks (R)	.358	–.148									.295
Loses temper	.915		*–.356*								.022
Steals	.646		.289								.498
Fights	.688		.339								.411
Lies, cheats	.729		.439								.276
Well behaved (R)	.385		*–.045*								.466
Internalizing behaviors											
Fearful others dislike me				.716	–.124						.471
Nervous				.725	.294						.388
Worry not as good as others				.765	–.251						.350
Worry about success				.724	–.191						.439
Generally concerned				.824	.447						.122
Worry too much				.723	.266						.406
Worry about future				.666	–.230						.505
Worry doing thinks right				.711	–.221						.445
Worry about past				.711	*.063*						.491
Feel depressed				.570		.677					.216
My life is a failure				.589		.625					.262
Shy				.544		.679					.244
Feel alone				.569		.708					.175
Often cries				.556		.740					.143
Feel sad				.716		.679					.471
Engagement											
Missed school							.403	.465			.622
Disrupt class							.431	.618			.432
Missed class							.295	.583			.573
Disrespect teacher							.449	.743			.246
Likes assignments							.571		.583		.334

Items	G-Ext λ	S-H/I λ	S-O/D λ	G-Int λ	S-Anx λ	S-Dep λ	G-Eng λ	S-BE λ	S-EE λ	S-CE λ	δ
Likes school							.594		.639		.239
Enjoys school							.516		.620		.349
Interested in schoolwork							.639		.453		.387
Do not want to stop working							.430		.310		.719
Happy to learn							.643		.347		.466
Plan schedule							.664			.394	.404
Identify important info.							.642			.536	.301
Try to solve problems							.652			.478	.347
Check understanding							.667			.487	.318
Link information							.680			.444	.341
Use strategies							.534			.474	.491
Ideas in my own words							.568			.501	.426
Composite reliability (ω)	.915	.742	.563	.952	.546	.929	.926	.756	.777	.807	

Note. G: Global factor from a bifactor solution; S: Specific factor from a bifactor solution; Ext: Externalizing behaviors; H/I: Hyperactivity/Inattention; O/D: Oppositional/Defiant; Int: Internalizing behaviors; Anx: Anxiety; Dep: Depression; Eng: Engagement; BE: Behavioral engagement; EE: Emotional engagement; CE: Cognitive engagement; ω: omega coefficient of composite reliability (McDonald, 1970).
Non-statistically significant results ($p \leq .05$) are marked in italics.

Table S4

Measurement and structural invariance tests across sex and immigration status in the secondary school sample.

Model	χ^2	*df*	CFI	TLI	RMSEA	RMSEA 90% CI	$\Delta\chi^2$	Δdf	ΔCFI	ΔTLI	ΔRMSEA
					Measurement models						
Bifactor-CFA Externalizing behaviors by sex											
1. Configural invariance	63.857*	44	.996	.992	.030	.010–.045	–	–	–	–	–
2. Weak invariance	76.706	63	.997	.996	.021	.000–.035	14.077	19	.001	.004	–.009
3. Strong invariance	81.016	69	.998	.997	.018	.000–.033	7.028	6	.001	.001	–.003
4. Strict invariance	85.116	79	.999	.999	.012	.000–.028	5.154	10	.001	.002	–.006
5. Latent variance invariance	100.004	83	.997	.996	.020	.000–.033	8.928	4	–.002	–.003	.008
6. Latent mean invariance	129.427*	87	.991	.991	.031	.019–.042	17.438*	4	–.006	–.005	.011
Bifactor-CFA Externalizing behaviors by immigration status											
1. Configural invariance	67.508*	44	.995	.990	.032	.015-.047	–	–	–	–	–
2. Weak invariance	80.475	63	.996	.995	.023	.000-.037	17.652	19	.001	.005	–.009
3. Strong invariance	84.325	68	.997	.996	.022	.000-.036	4.939	5	.001	.001	–.001
4. Strict invariance	121.525*	78	.991	.990	.033	.021-.044	39.057*	10	–.006	–.006	.011
5. Latent variance invariance	128.668*	82	.990	.990	.033	.022-.044	8.310	4	–.001	.000	.000
6. Latent mean invariance	360.572*	86	.944	.941	.079	.071-.087	123.855*	4	–.046	–.049	.046
6. Latent mean invariance–partial	143.460*	84	.988	.987	.037	.026–.047	15.137*	2	–.002	–.003	.004
Bifactor-CFA Internalizing behaviors by sex											
1. Configural invariance	455.030*	150	.986	.981	.063	.057–.070	–	–	–	–	–
2. Weak invariance	543.013*	177	.983	.980	.064	.058–.070	102.517*	27	–.003	–.001	.001
3. Strong invariance	532.915*	207	.985	.985	.056	.050–.062	34.894	30	.002	.005	–.008
4. Strict invariance	466.011*	222	.989	.990	.047	.041–.053	21.059	15	.004	.005	–.009
5. Latent variance invariance	376.900*	225	.993	.994	.037	.030–.043	4.672	3	.004	.004	–.010
6. Latent mean invariance	900.104*	228	.970	.972	.076	.071–.082	138.749*	3	–.023	–.022	.039
6. Latent mean invariance–partial	359.425*	227	.994	.994	.034	.027–.040	89.237*	2	.001	.000	–.003
Bifactor-CFA Internalizing behaviors by immigration status											
1. Configural invariance	921.418*	150	.969	.957	.101	.095–.107	–	–	–	–	–
2. Weak invariance	869.122*	177	.984	.982	.066	.060–.072	306.063*	27	.015	.025	–.035
3. Strong invariance	471.370*	207	.990	.989	.050	.044–.056	342.600*	30	.006	.007	–.016
4. Strict invariance	471.580*	222	.990	.991	.047	.041–.053	44.005*	15	.000	.002	–.003
5. Latent variance invariance	361.759*	225	.995	.995	.035	.028–.041	3.030	3	.005	.004	–.012
6. Latent mean invariance	404.300*	228	.993	.994	.039	.033–.045	16.635*	3	–.002	–.001	.004
Bifactor-CFA Engagement by sex											
1. Configural invariance	627.648*	204	.975	.966	.064	.058–.069	–	–	–	–	–
2. Weak invariance	592.639*	239	.979	.976	.054	.048–.059	57.178*	35	.004	.010	–.010
3. Strong invariance	594.745*	282	.981	.982	.047	.041–.052	52.659	43	.002	.006	–.007
4. Strict invariance	606.346*	299	.982	.983	.045	.040–.050	45.826*	17	.001	.001	–.002

Model	χ^2	*df*	CFI	TLI	RMSEA	RMSEA 90% CI	$\Delta\chi^2$	Δdf	ΔCFI	ΔTLI	ΔRMSEA
5. Latent variance invariance	468.320*	303	.990	.991	.033	.027–.038	2.590	4	.008	.008	–.012
6. Latent mean invariance	567.188*	307	.984	.986	.041	.035–.046	36.095*	4	–.006	–.005	.008
Bifactor-CFA Engagement by immigration status											
1. Configural invariance	599.850*	204	.976	.968	.061	.056–.067	–	–	–	–	–
2. Weak invariance	553.743*	239	.981	.979	.051	.045–.056	46.010	35	.005	.011	–.010
3. Strong invariance	598.347*	282	.981	.982	.047	.042–.052	81.926*	43	.000	.003	–.004
4. Strict invariance	564.261*	299	.984	.986	.042	.036–.047	19.861	17	.003	.004	–.005
5. Latent variance invariance	443.699*	303	.992	.992	.030	.024–.036	2.965	4	.008	.006	–.012
6. Latent mean invariance	454.469*	307	.991	.992	.031	.024–.036	9.676*	4	–.001	.000	.001
Predictive (Structural) Model											
Invariance by sex											
1. Configural invariance	23.228	14	.993	.931	.036	.000–.061	–	–	–	–	–
2. Regression invariance	109.622*	73	.971	.947	.032	.018–.043	86.505*	59	–.022	.016	–.004
2. Regression invariance–partial	88.852	71	.986	.974	.022	.000–.036	65.827	57	–.007	.043	–.014
3. Intercept invariance	127.470*	75	.958	.927	.037	.026–.048	36.789*	4	–.028	–.047	.015
3. Intercept invariance–partial	92.286	73	.985	.972	.023	.000–.036	3.425	2	–.001	–.002	.001
4. Residuals invariance	99.241*	77	.982	.970	.024	.004–.037	6.897	4	–.003	–.002	.001
Invariance by immigration status											
1. Configural invariance	37.907*	14	.981	.824	.058	.036–.081	–	–	–	–	–
2. Regression invariance	100.965*	72	.977	.958	.028	.013–.040	63.239	58	–.004	.134	–.030
3. Intercept invariance	119.340*	76	.966	.941	.034	.012–.045	18.643*	4	–.011	–.017	.006
3. Intercept invariance–partial	106.740*	75	.975	.956	.029	.015–.041	5.835	3	–.002	–.002	.001
4. Residuals invariance	111.895*	79	.974	.957	.029	.015–.040	5.201	4	–.001	.001	.001

Note. χ^2: Chi square test of model fit and associated degrees of freedom (*df*); CFI: Comparative Fit Index; TLI: Tucker-Lewis Index; RMSEA: Root Mean Square Error of Approximation and 90% Confidence Interval (CI); Δ: Change according to the previous retained model; $\Delta\chi^2$: Chi square difference test calculated with the Mplus DIFFTEST option for the measurement invariance model and with the Satorra-Bentler correction for the predictive (structural) models.
*$p < .05$.