Factorial Invariance of the Academic Amotivation Inventory (AAI) Across Gender and Grade in a Sample of Canadian High School Students
Isabelle Green-Demers, Lisa Legault, Daniel Pelletier and Luc G. Pelletier
Educational and Psychological Measurement 2008; 68; 862 originally published online Mar 4, 2008;
DOI: 10.1177/0013164407313368

The online version of this article can be found at:
http://epm.sagepub.com/cgi/content/abstract/68/5/862

Published by:
SAGE
http://www.sagepublications.com

Additional services and information for Educational and Psychological Measurement can be found at:

Email Alerts: http://epm.sagepub.com/cgi/alerts

Subscriptions: http://epm.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations http://epm.sagepub.com/cgi/content/refs/68/5/862
Motivation deficits are common in high school and constitute a significant problem for both students and teachers. The Academic Amotivation Inventory (AAI) was developed to measure the multidimensional nature of the academic amotivation construct (Legault, Green-Demers, & Pelletier, 2006). The present project further examined the consistency of the metric properties of AAI scores by testing their factorial structure for invariance across gender and grade (2 [genders] × 5 [grades] = 10 [groups]) in a sample of 3,417 high school students. Factorial invariance of latent means was also examined as a complementary substantive objective. Configural, metric, and scalar invariance were successfully substantiated across all 10 groups. Results revealed well-fitting models for each group. Moreover, constraining factor loadings and intercepts had no meaningful impact on model fit. Findings are discussed in terms of an increased conceptual and psychometric understanding of scholastic motivational problems.

Keywords: school amotivation; factorial invariance; grade; gender; high school

There is a wealth of available information pertaining to academic motivation. It is widely recognized that high student motivation is beneficial; it is related to a host of positive outcomes (for reviews please see Chouinard, 2001; Pintrich, 2003; Reeve, 2002; Ryan & Deci, 1999; and Educational Psychologist, 1991, Vol. 4, for a complete number on this topic). For instance, academic motivation has been associated with greater cognitive flexibility, conceptual understanding, and active information processing (Grolnick & Ryan, 1987); higher learning self-regulation...
(Zimmerman & Pons, 1990); increased time spent studying; lower tardiness and absenteeism (Yelle, Green-Demers, & Pelletier, 2005); better academic performance and self-concept (Deci, Vallerand, Pelletier, & Ryan, 1991; Reeve, Bolt, & Cai, 1999); and more ambitious plans for postsecondary education (Green-Demers & Pelletier, 2003; Vallerand, Blais, Brière, & Pelletier, 1989). Self-determined academic motivation has also been found to protect against high school dropout (Hardre & Reeve, 2003; Vallerand, Fortier, & Guay, 1997).

However, the literature pertaining to high school dropout makes it clear that a substantial proportion of high school students lack motivation in the classroom. Indeed, a high number of young Canadians and Americans abandon high school every year (Snyder & Hoffman, 2002; Statistics Canada, 2002). Evidently, this kind of trend has enormous social, psychological, and economical ramifications. For the student, high school dropout can result in decreased quality of life, both present and future (Laflueur, 1992). It can lead to restrictions in employment, dependence on social compensation, and even criminal behavior (Garnier, Stein, & Jacob, 1997; Newcomb et al., 2002). A deeper understanding of academic amotivation may prove very useful in comprehending and preventing dropout. Indeed, evidence suggests that motivational deficits are important factors in students’ quitting high school. Janosz and his colleagues, for instance, found that academic disengagement was an excellent predictor of dropout (Janosz, 2000; Janosz, Leblanc, Boulerice, & Tremblay, 1997). Yet by contrast to the vast extant theoretical and empirical documentation on academic motivation, little attention has been paid to the reasons why some students are inclined to neglect their studies. It has been argued that a better understanding of the reasons why students lack academic motivation is a contemporary issue of critical importance (Hidi & Harackiewicz, 2000).

In an attempt to examine this issue, we developed a taxonomy that comprises various forms of academic motivation deficit (Legault, Green-Demers, & Pelletier, 2006). An overview of the theoretical underpinnings of our approach is offered below. Information pertaining to the multidimensional nature of academic amotivation is presented thereafter.

**Conceptualizing Amotivation**

According to Self-Determination Theory (Deci & Ryan, 1985, 2002), there are three broad categories of behavior regulation: intrinsic motivation (pleasure motives);
extrinsic motivation (instrumental motives, which can be further differentiated into several subcategories on the basis of their level of autonomy); and amotivation (an absence of motivation). The various intrinsic and extrinsic motives represent distinct types of behavior regulation. By contrast, amotivation, the central tenet of the present study, is defined as a form of dysregulation that is characterized by a dissociation between behavior and outcome. Such behaviors are thus perceived as devoid of meaning. They can be performed for reasons unknown (e.g., out of habit), in a dispirited manner, or eschewed altogether. Amotivated individuals feel disintegrated or detached from their actions, invest little effort and energy in their effectuation, and are likely to give up entirely. Amotivation is experienced as a lack of agency and control that has been compared to learned helplessness (Abramson, Seligman, & Teasdale, 1978).

There is a dearth of research pertaining to amotivation. A mere handful of studies has addressed the correlates of amotivation in sports and education. Specifically, amotivation has been related to attrition among competitive swimmers (Pelletier, Fortier, Vallerand, & Brière, 2001) and handball players (Sarrazin, Vallerand, Guillet, Pelletier, & Cury, 2002) and has also been associated with boredom, low attendance, and low involvement in physical education (Ntoumanis, Pensgaard, Martin, & Pipe, 2004). Moreover, in the academic domain, amotivation has been correlated with boredom and poor concentration in class (Vallerand et al., 1993), higher perceived school-related stress, poor psychosocial adjustment to university (Baker, 2004), and high school dropout (Vallerand & Bissonnette, 1992; Vallerand et al., 1997). In all of the studies described in this paragraph, amotivation was conceptualized as a unidimensional construct and was measured by means of a single subscale that was included in multidimensional motivation self-report questionnaires.

**Conceptualizing Amotivation as a Multidimensional Construct**

Although Self-Determination Theory’s (Deci & Ryan, 1985, 2002) traditional view of amotivation as a unidimensional concept is of seminal importance, it has been suggested that motivational deficits are manifold and that amotivation is a structurally complex phenomenon that can be conceptualized as a multifaceted construct. For instance, in a prior study that pertained to the construction and validation of an instrument devised to evaluate the lack of motivation to perform ecological behaviors (e.g., recycling, conserving energy), amotivation was found to stem from four different classes of reasons: helplessness beliefs as well as deficiencies in strategy beliefs, ability beliefs, and effort beliefs (Pelletier, Dion, Tuson, & Green-Demers, 1999). The four dimensions of environmental amotivation proposed by Pelletier and his colleagues were supported by exploratory and confirmatory factor analyses. This inspired us to theorize that the debilitating process of academic inertia is a multifarious phenomenon that may express itself in different forms (Legault et al., 2006). Specifically, we proposed that students were liable to lack motivation in school for...
different kinds of reasons. From the four dimensions proposed by Pelletier and his colleagues, two were retained as they are relevant in the academic domain: deficits in ability beliefs and effort beliefs. Two additional constructs were included: insufficient academic values and unappealing characteristics of school tasks.

**Deficient ability beliefs.** This form of amotivation is said to occur when students attribute their academic difficulties to low perceived competence (Wigfield, 1988), when they hold low self-efficacy expectancies about schoolwork (Skinner, Wellborn, & Connell, 1990), and when their academic ability self-concept is poor or ill defined (Eccles et al., 1993).

**Deficient effort beliefs.** Amotivation may also stem from a lack of desire or capacity to exert the effort required by academic behaviors. Skinner and her colleagues (1990) demonstrated that ability and effort beliefs were both necessary antecedents of school performance. Academic detachment resulting from a lack of ability or desire to exert effort has also been noted by several authors (e.g., Chouinard, 2001; Eccles et al., 1993; Patrick, Skinner, & Connell, 1993).

**Insufficient academic values.** People need good reasons to perform effortful behaviors. If academic tasks are unimportant to a student, amotivation may ensue. Recent important articles pertaining to Self-Determination Theory have included a lack of value as part of the definition of amotivation (Ryan, 1995; Ryan & Deci, 1999, 2000). Many researchers have argued that values are key to understanding academic behaviors (e.g., Bigelow & Zhou, 2001; Pintrich & De Groot, 1990; Wigfield & Eccles, 1992, 1994).

**Unappealing characteristics of school tasks.** When a task is perceived as void of interesting or stimulating qualities, that is, if it is seen as boring and tedious, it is unlikely to engage students (Ainley, Hidi, & Berndoff, 2002) and may constitute a factor in amotivation.

To summarize, we proposed a taxonomy of academic amotivation that comprises four types of motivational deficits (i.e., insufficient ability beliefs, effort beliefs, academic values, and unappealing task characteristics). This undertaking led to the construction of the Academic Amotivation Inventory (AAI; Legault et al., 2006).

**Development of the AAI and Initial Psychometric Findings**

The AAI is an instrument designed to measure the four academic amotivation dimensions described above in high school students. Three studies reported in an article published in the *Journal of Educational Psychology* (Legault et al., 2006)
were initially performed to ascertain the psychometric properties of the AAI. In the course of this series of studies, the factorial structure of the AAI was supported by the results of exploratory (Study 1, \( N = 351 \)) and confirmatory factor analyses (Study 2, \( N = 349 \); Study 3, \( N = 741 \)). The construct validity of the scores of the AAI was documented by associations with relevant behavioral and psychological academic variables (Study 2), as well as by means of structural equation models that examined relationships between social support (from teachers, parents, and friends), academic amotivation, and academic outcomes (Study 3). The discriminant validity of amotivation dimensions with respect to the motivation subtypes of Self-Determination Theory (Deci & Ryan, 1985, 2002) was also supported by the results of Studies 2 and 3. The internal consistency (coefficient alpha) of the subscale scores for the final version of the AAI was found to be very satisfactory (i.e., \( .84 < \alpha < .86 \) in Study 2 and \( .81 < \alpha < .86 \) in Study 3). Finally, metric multigroup invariance of confirmatory factor analysis models across gender was successfully documented in Study 3.

**Objectives**

The main goal of this project was to further examine the consistency of the factorial structure of the AAI, using an independent and much larger sample of high school students. Specifically, the factorial structure of the AAI was tested using confirmatory models for boys and girls of all high school levels. Results were compared using multigroup invariance testing of factor loadings and intercepts across gender and grade (2 [genders] \( \times \) 5 [grades] = 10 [groups]). The secondary objective of this study was to assess and compare latent means across gender and grade. This substantive endeavor is intended as a complement to the more central psychometric analyses. Assessing measurement and latent mean invariance across gender and grade is meaningful because these variables have been shown to influence motivational dynamics (please refer to paragraph below). From a psychometric point of view, it is important to ascertain whether the AAI performs consistently for boys and girls of all grades. From a substantive point of view, it is interesting to assess gender and grade effects on latent means to expand our understanding of their impact on multidimensional motivation deficits.

**Gender, Grade, and Amotivation**

The results of past studies pertaining to the influence of gender on educational outcomes are typically more favorable for girls than boys (please refer to Martin, 2004, for a literature review on this topic). Motivational studies yielded comparable results. Girls have been found to have higher self-determined motivation than boys.
Conversely, external motivation and amotivation (measured as a unidimensional construct) are typically higher in boys than girls (Grouzet, Otis, & Pelletier, 2006; Ratelle, Guay, Larose, & Sénécal, 1994; Vallerand et al., 1997). These effects are of modest magnitude but are consistent across studies. As for the effect of grade, the documentation indicates that motivation declines over time (Eccles et al., 1993; Grouzet et al., 2006; Harter, Whitesell, & Kowalski, 1992; Otis, Grouzet, & Pelletier, 2005; Peetsma, Hascher, & van der Veen Ewoud Roede, 2005).

It was difficult to make precise predictions regarding possible latent mean differences in the present study because it is the first time that the influence of gender and grade is evaluated concurrently for several amotivation dimensions. The magnitude of latent means for gender and grade was therefore examined in an exploratory vein. No specific hypotheses were formulated, although it was expected that differences in latent means would indeed occur, at least for some amotivation dimensions, and that those results would follow the trends in the literature.

**Method**

**Participants and Procedure**

Data were collected from 3,417 francophone high school students in the Ottawa–Gatineau region (Canada). Students were aged 12 to 18 years, with a mean age of 14 years ($SD = 2.38$). Their self-reported grade point average was 76.19% ($SD = 9.30$%). The sample included 1,622 boys and 1,795 girls (the specific number of boys and girls in Grades 8 to 12 can be found in Table 1). Students completed the questionnaire at school during class time.

**Instrument**

The principal measure of interest, L’Inventaire de Motivation Académique (i.e., the AAI), comprises four subscales (four items per subscale) devised to ascertain the four proposed dimensions of academic amotivation: deficits in ability beliefs, deficits in effort beliefs, lack of academic values, and unappealing characteristics of academic tasks (Legault et al., 2006). Items are rated on a 7-point scale ($1 =$ corresponds exactly; $4 =$ corresponds moderately; $7 =$ corresponds exactly) and are presented in random order.

**Analyses**

*Measurement invariance testing.* Confirmatory factor analyses with maximum likelihood estimation were performed using the EQS program (v. 6.1; Bentler, 2006). Measurement invariance implies that the indicators of a construct reflect equivalent
domain representations, and is defined as the mathematical equality of measurement parameters (Little, 1997). Invariance testing was performed in three steps, using increasingly stringent nested models. Configural invariance was first assessed (i.e., the tenability of the four-factor structure of academic amotivation was examined in each group). Metric invariance (equivalency of factor loadings; Meredith, 1993, “weak invariance”) and scalar invariance (equivalency of factor intercepts; Meredith, 1993, “strong invariance”) were evaluated thereafter in sequential steps.

### Table 1

**Internal Consistency (Cronbach’s $\alpha$) of Subscale Scores in Gender and Grade Subgroups**

<table>
<thead>
<tr>
<th>Subsample</th>
<th>Ability Beliefs</th>
<th>Effort Beliefs</th>
<th>Value of School</th>
<th>Task Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys ($n = 477$)</td>
<td>.87 (.85 $&lt; \alpha &lt; .89$)</td>
<td>.82 (.79 $&lt; \alpha &lt; .85$)</td>
<td>.85 (.82 $&lt; \alpha &lt; .87$)</td>
<td>.87 (.85 $&lt; \alpha &lt; .89$)</td>
</tr>
<tr>
<td>Girls ($n = 492$)</td>
<td>.89 (.87 $&lt; \alpha &lt; .91$)</td>
<td>.83 (.81 $&lt; \alpha &lt; .86$)</td>
<td>.85 (.83 $&lt; \alpha &lt; .87$)</td>
<td>.88 (.86 $&lt; \alpha &lt; .89$)</td>
</tr>
<tr>
<td>Grade 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys ($n = 364$)</td>
<td>.90 (.88 $&lt; \alpha &lt; .92$)</td>
<td>.83 (.80 $&lt; \alpha &lt; .86$)</td>
<td>.87 (.85 $&lt; \alpha &lt; .89$)</td>
<td>.89 (.87 $&lt; \alpha &lt; .91$)</td>
</tr>
<tr>
<td>Girls ($n = 401$)</td>
<td>.90 (.88 $&lt; \alpha &lt; .91$)</td>
<td>.85 (.82 $&lt; \alpha &lt; .87$)</td>
<td>.89 (.87 $&lt; \alpha &lt; .91$)</td>
<td>.88 (.86 $&lt; \alpha &lt; .90$)</td>
</tr>
<tr>
<td>Grade 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys ($n = 322$)</td>
<td>.87 (.84 $&lt; \alpha &lt; .89$)</td>
<td>.77 (.73 $&lt; \alpha &lt; .81$)</td>
<td>.86 (.84 $&lt; \alpha &lt; .89$)</td>
<td>.85 (.82 $&lt; \alpha &lt; .88$)</td>
</tr>
<tr>
<td>Girls ($n = 348$)</td>
<td>.85 (.82 $&lt; \alpha &lt; .87$)</td>
<td>.79 (.76 $&lt; \alpha &lt; .83$)</td>
<td>.83 (.79 $&lt; \alpha &lt; .85$)</td>
<td>.85 (.83 $&lt; \alpha &lt; .88$)</td>
</tr>
<tr>
<td>Grade 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys ($n = 277$)</td>
<td>.88 (.86 $&lt; \alpha &lt; .91$)</td>
<td>.77 (.72 $&lt; \alpha &lt; .81$)</td>
<td>.87 (.84 $&lt; \alpha &lt; .89$)</td>
<td>.85 (.82 $&lt; \alpha &lt; .88$)</td>
</tr>
<tr>
<td>Girls ($n = 330$)</td>
<td>.87 (.85 $&lt; \alpha &lt; .89$)</td>
<td>.78 (.74 $&lt; \alpha &lt; .82$)</td>
<td>.81 (.77 $&lt; \alpha &lt; .84$)</td>
<td>.86 (.83 $&lt; \alpha &lt; .88$)</td>
</tr>
<tr>
<td>Grade 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys ($n = 182$)</td>
<td>.87 (.84 $&lt; \alpha &lt; .90$)</td>
<td>.76 (.70 $&lt; \alpha &lt; .81$)</td>
<td>.86 (.82 $&lt; \alpha &lt; .89$)</td>
<td>.84 (.79 $&lt; \alpha &lt; .87$)</td>
</tr>
<tr>
<td>Girls ($n = 224$)</td>
<td>.84 (.81 $&lt; \alpha &lt; .87$)</td>
<td>.82 (.78 $&lt; \alpha &lt; .85$)</td>
<td>.84 (.81 $&lt; \alpha &lt; .88$)</td>
<td>.88 (.85 $&lt; \alpha &lt; .90$)</td>
</tr>
<tr>
<td>Total ($N = 3,417$)</td>
<td>.88 (.87 $&lt; \alpha &lt; .89$)</td>
<td>.82 (.81 $&lt; \alpha &lt; .83$)</td>
<td>.87 (.86 $&lt; \alpha &lt; .87$)</td>
<td>.87 (.86 $&lt; \alpha &lt; .88$)</td>
</tr>
</tbody>
</table>

Note: Values in parentheses represent 95% confidence intervals.

**Latent mean invariance testing.** According to the means and covariance structures approach (MACS; Little, 1997), construct comparability across groups is assessed by testing for the equality of measurement parameters (e.g., invariance of factor loadings and intercepts, as described above). The measurement level is distinguished from the latent construct level and is designed to evaluate possible group differences (e.g., discrepancies in latent means). Measurement invariance and constraints must be evaluated and implemented before latent construct analyses are performed. It must be underscored that whereas measurement invariance results (e.g., strong invariance) bear psychometric implications, this is not the case for latent construct results. Differences in latent means, for instance, are not a reflection of the consistency of scores from an instrument. They can be construed as the latent construct counterpart of simple analyses designed to evaluate observed mean.
differences (e.g., ANOVAs). Their purpose is to document similarities and differences in the magnitude of factor scores across groups. Such findings are of substantive rather than psychometric interest.

**Assessment of model fit.** In the present study, for all analyses, the degree of model fit was assessed from several angles using several criteria (i.e., the Satorra-Bentler scaled statistic \([S-B\chi^2]\); Satorra & Bentler, 1988), the comparative fit index (CFI; Bentler, 1990), the root mean square error of approximation (RMSEA; Steiger, 1989), and the standardized root mean square residual (SRMR; Jöreskog & Sörbom, 1993). The \(S-B\chi^2\) was reported instead of the \(\chi^2\), as it provides an adjustment that protects against potential deviations from the assumption of multivariate normality. Please note that the CFI and RMSEA values reported herein are based on the \(S-B\chi^2\) (Bentler, 2006). The CFI is an incremental fit index (Bollen, 1989) that is not unduly influenced by sample size (Marsh, Balla, & McDonald, 1988). Values above .90 have traditionally been considered acceptable (Cheung & Rensvold, 2002; Kline, 1998; McDonald & Ho, 2002). However, Hu and Bentler (1999) recommended that higher standards be entertained (i.e., cutoff values close to .95; Shumacker & Lomax, 2004). The RMSEA evaluates the estimated discrepancy per degree of freedom between the population covariance matrix and the model, whereas the SRMR represents the standardized average of covariance residuals. The RMSEA and SRMR are sensitive to model parsimony. Values smaller than .08 suggest a reasonable fit (Bentler, 2006; Browne & Cudeck, 1993).

Differences between nested models were evaluated by computing the degree of change between Satorra-Bentler scaled statistics \((\Delta S-B\chi^2)\) and CFI values \((\Delta CFI)\). The degree of change in \(S-B\chi^2\) was calculated following the strategy recommended by Satorra and Bentler (2001). Changes in CFI values simply represent the difference between the CFI of the baseline and constrained models. Cheung and Rensvold (2002) note that many changes in goodness-of-fit indices (e.g., \(\Delta CFI\)) are superior to changes in chi-square as tests of invariance because they are not affected by sample size. Moreover, these authors argue that \(\Delta CFI\) is one of the few difference statistics that is independent of both model complexity and sample size and is not correlated with overall fit measures. These authors conclude that a value of \(\Delta CFI\) smaller than or equal to −0.01 indicates that the null hypothesis of invariance should not be rejected.

**Results**

**Preliminary Analyses**

*Reliability of subscale scores.* The internal consistency (Cronbach’s \(\alpha\)) of test scores was calculated globally for each subscale and individually for each gender
and grade subgroup. Confidence intervals (95%) were also evaluated using the method recommended by Fan and Thompson (2001). Reporting confidence intervals is increasingly recommended because reliability affects Type I and II error rates as well as effect sizes (Iacobucci & Duhachek, 2003; Onwuegbuzie & Daniel, 2002). Results are reported in Table 1. The overall internal consistency of test scores for each subscale was comparable to that obtained in the original validation studies (Legault et al., 2006). Overall coefficients (and their confidence limits) were also entirely above the .80 cutoff that is typically considered acceptable for general research purposes (Henson, 2001; Loo, 2001; Nunnally & Bernstein, 1994). Internal consistency within each of the 10 gender and grade subgroups examined herein was also largely acceptable. Of the 40 coefficients, 35 (88%) were above .80. The 5 coefficients that were somewhat weaker were above .75. Confidence intervals for all subgroups included .80. The lowest among lower-bound confidence limits was .70.

Intraclass correlations. Data that are collected in schools are often structured hierarchically (i.e., students within schools). If such is the case and if the variance of school membership is left unaccounted for, there may be systematic variance in error residuals caused by the confounding influence of this variable. In addition to the loss of information that results from omitting a meaningful variable, this violation of the assumption of independence of error can result in a steep inflation of Type 1 error rates (Barcikowski, 1981; Heck & Thomas, 2000). Muthén (1994) recommended generating intraclass correlations (ICCs) to ascertain whether the hierarchical structure of a data set should be modeled using multilevel analyses. ICCs represent the proportion of sample variance that is explained by the grouping variable (e.g., school membership; Hox, 2002; Tabachnick & Fidell, 2007). The influence of the grouping variable should be modeled when ICCs are superior to .10 (Muthén, 1997). ICCs were computed for all 16 items of the AAI. Values varied from .001 to .019 (M = .009). These values are close to zero and well below the .10 cutoff cited above (Muthén, 1997). This suggests that between-school variance is trivial in our sample and that it is unnecessary to model its influence.

Multigroup Invariance Testing

As previously noted, measurement invariance testing was performed in three steps: configural invariance, metric invariance, and scalar invariance (Little, 1997; Meredith, 1993). Configural invariance results are reported in Table 2, whereas metric and scalar invariance results are displayed in Table 3.

Configural invariance was first ascertained by testing the four-factor structure of the AAI separately for all gender and grade groups, without any constraints. These analyses were specified as typical measurement models wherein target loadings, item uniqueness values, and factor variances and covariances were freely estimated. For identification purposes, the first loading of each factor was fixed to 1.00. Results
revealed that, according to all relevant and substantively meaningful fit indices discussed herein, model fit was satisfactory. No post hoc model respecifications were required for any of the 10 groups. Moreover, for each group, all estimated parameters (factor loadings, factor variances and covariances, and item uniqueness values) were of satisfactory magnitude and were statistically significant at the $p < .01$ level.

Metric invariance was assessed next by constraining factor loadings to be equivalent in all 10 groups. Overall indices revealed that the fit of the constrained model

---

Table 2
Confirmatory Factor Analyses: Model Fit for Each Gender and Grade Subgroup

<table>
<thead>
<tr>
<th>Subsample</th>
<th>S-B$\chi^2$</th>
<th>$df$</th>
<th>CFI</th>
<th>SRMR</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grade 8</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys ($n = 477$)</td>
<td>306.586**</td>
<td>98</td>
<td>.939</td>
<td>.059</td>
<td>.067 (.058; .075)</td>
</tr>
<tr>
<td>Girls ($n = 492$)</td>
<td>264.072**</td>
<td>98</td>
<td>.947</td>
<td>.051</td>
<td>.059 (.050; .067)</td>
</tr>
<tr>
<td><strong>Grade 9</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys ($n = 364$)</td>
<td>275.187**</td>
<td>98</td>
<td>.950</td>
<td>.058</td>
<td>.071 (.061; .080)</td>
</tr>
<tr>
<td>Girls ($n = 401$)</td>
<td>212.663**</td>
<td>98</td>
<td>.966</td>
<td>.047</td>
<td>.054 (.044; .064)</td>
</tr>
<tr>
<td><strong>Grade 10</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys ($n = 322$)</td>
<td>171.448**</td>
<td>98</td>
<td>.970</td>
<td>.051</td>
<td>.048 (.036; .060)</td>
</tr>
<tr>
<td>Girls ($n = 348$)</td>
<td>214.705**</td>
<td>98</td>
<td>.943</td>
<td>.067</td>
<td>.059 (.048; .069)</td>
</tr>
<tr>
<td><strong>Grade 11</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys ($n = 277$)</td>
<td>188.710**</td>
<td>98</td>
<td>.954</td>
<td>.063</td>
<td>.058 (.045; .070)</td>
</tr>
<tr>
<td>Girls ($n = 330$)</td>
<td>270.118**</td>
<td>98</td>
<td>.924</td>
<td>.068</td>
<td>.073 (.063; .083)</td>
</tr>
<tr>
<td><strong>Grade 12</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys ($n = 182$)</td>
<td>174.879**</td>
<td>98</td>
<td>.942</td>
<td>.072</td>
<td>.066 (.050; .081)</td>
</tr>
<tr>
<td>Girls ($n = 224$)</td>
<td>193.189**</td>
<td>98</td>
<td>.943</td>
<td>.067</td>
<td>.066 (.052; .079)</td>
</tr>
</tbody>
</table>

Note: CFI = comparative fit index; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation. Values in parentheses represent 90% confidence intervals.

**$p < .001$.**

---

Table 3
Invariance Testing Across Gender and Grade: Fit Indices of Nested Models

<table>
<thead>
<tr>
<th>Models</th>
<th>Overall Fit Indices</th>
<th>Comparative Fit Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S-B$\chi^2$</td>
<td>$df$</td>
</tr>
<tr>
<td>Baseline</td>
<td>2307.854**</td>
<td>98</td>
</tr>
<tr>
<td>Metric invariance</td>
<td>2453.389**</td>
<td>1076</td>
</tr>
<tr>
<td>Scalar invariance</td>
<td>2917.671**</td>
<td>1204</td>
</tr>
</tbody>
</table>

Note: CFI = comparative fit index; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation. Values in parentheses represent 90% confidence intervals.

*$p < .05$. **$p < .001$.**
was satisfactory. According to strictly statistical criteria (ΔS-Bχ²), the difference in adjustment between the baseline and the constrained model was statistically significant. However, the ΔCFI was inferior to −.01, thereby suggesting that there are no meaningful discrepancies between groups (Cheung & Rensvold, 2002).

Scalar invariance was evaluated as the last step of measurement invariance testing, by constraining factor loadings and intercepts to be equivalent across groups. The overall model fit was satisfactory. As in the case of metric invariance, the ΔS-Bχ² was statistically significant. Yet the ΔCFI was inferior to −.01, indicating that constraining factor loadings and intercepts across gender and grade had little substantive influence on model fit.

Lastly, to complement measurement invariance analyses, possible differences in average latent factor scores were evaluated across gender and grade. As measurement invariance is a preliminary step to latent mean invariance testing, model specifications included factor loadings and intercept constraints. In addition to the model specifications and constraints that were required for scalar invariance, the current analyses required the estimation of factor means. Please note that latent factor intercepts (means) have an arbitrary origin (Bentler, 2006). Latent mean differences are scale free and a comparison standard is established by fixing them to zero in one group (Byrne, 2006). In the current analysis, the means for Grade 8 boys were set to zero.

To fully document which of the 40 factor means differed from one another, four sequential analyses were performed. In the first of these analyses, additional constraints were required because the number of estimated means (40) was too high to achieve model identification by simply fixing the means in one group to zero. For each amotivation dimension, all estimated means were fixed to be equal for boys and girls within grades, thereby diminishing the number of estimated means by half. In the second analysis, all factor means that were not statistically significantly different from those of Grade 8 boys were also fixed to zero, which enabled us to free up the previously imposed gender constraints for which the Lagrange multiplier statistic was significant. In the third run, the relative magnitude of estimated means was subjectively evaluated and the ones that appeared similar were constrained to equality. In the fourth and final run, the constraints that did not hold (according to the Lagrange multiplier statistic) were released.

Results indicated an acceptable model fit, S-Bχ²(df = 1200) = 2765.914, p < .001; CFI = .948; SRMR = .071; RMSEA = .062; CI.90 = .060 < RMSEA < .066. Latent means are displayed in Table 4. Please recall that their magnitude is relative and is expressed as the difference from the average latent factor scores of Grade 8 boys. No gender or grade effects were detected for capacity beliefs. For this dimension, latent means were invariant across all 10 groups. The only amotivation construct for which gender effects were detected was effort beliefs. Within each grade, boys scored higher than girls. Grade 8 girls scored lower than Grade 8 boys. In Grades 9 to 12, girls’ effort beliefs were not significantly different from those of Grade 8 boys. However, for boys, amotivation due to effort beliefs increased...
over grades. It was higher in Grade 9 than in Grade 8, and higher still in Grades 10 to 12. Differences between grades were also identified for school value and task characteristic dimensions. For both of these constructs and for both genders, amotivation was higher in Grade 9 than in Grade 8 and higher yet in Grades 10 to 12. Comparing across dimensions, the magnitude of amotivation due to lack of school values in boys and girls (Grades 9 to 12) and amotivation due to lack of effort beliefs in boys (Grades 9 to 12) was higher than amotivation due to capacity beliefs. In turn, amotivation resulting from unappealing task characteristics in boys and girls (Grades 9 to 12) was higher than all other forms of amotivation. Effect sizes for latent means were computed for each factor using Hancock’s (2001) method.

Table 4
Latent Means Across Gender and Grade

<table>
<thead>
<tr>
<th></th>
<th>Grade 8</th>
<th>Grade 9</th>
<th>Grade 10</th>
<th>Grade 11</th>
<th>Grade 12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Boys</td>
<td>Girls</td>
<td>Boys</td>
<td>Girls</td>
<td>Boys</td>
</tr>
<tr>
<td>Ability beliefs</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>Value of school</td>
<td>.00</td>
<td>.00</td>
<td>.29</td>
<td>.29</td>
<td>.47</td>
</tr>
<tr>
<td>Effort beliefs</td>
<td>.00</td>
<td>−.17</td>
<td>.29</td>
<td>.00</td>
<td>.47</td>
</tr>
<tr>
<td>Task characteristics</td>
<td>.00</td>
<td>.00</td>
<td>.47</td>
<td>.47</td>
<td>.89</td>
</tr>
</tbody>
</table>

Note: For statistical identification purposes, latent means were fixed to 0 for Grade 8 boys. Latent means of different magnitude are statistically different from one another (p < .05). Effect sizes for each factor (f) were computed using Hancock’s (2001) method.

Discussion

The objective of the present study was to examine the functional equivalency of the AAI in high school students. Multigroup invariance testing across gender and grade revealed that the four-factor structure of the AAI was a satisfactory representation of motivational deficits for boys and girls of Grades 8 to 12. That is, configural invariance was obtained across all 10 gender and grade groups. Moreover, overall fit indices revealed that multigroup models comprising factor loadings and intercepts constraints displayed a satisfactory adjustment to sample covariance. Statistical discrepancies, as noted by the ΔS-Bχ², were observed between the baseline model and the model with factor loadings constrained as well as for the model with factor loadings and intercepts constrained. However, ΔCFI values substantiated the metric and scalar invariance of the AAI.

The discrepancy between ΔS-Bχ² and ΔCFI results can be readily attributed to the characteristics of our analyses. The χ² statistic is notoriously oversensitive to
sample size (e.g., Hu & Bentler, 1995) and to model complexity (Bollen, 1989). The present study comprised a relatively large sample. Invariance of factor loadings and/or intercepts across 10 groups also qualifies as a very complex covariance structure analysis. In this context, it is not surprising that significant ΔS-Bχ² differences were detected, and ΔCFI values were therefore considered more pertinent. As Cheung and Rensvold (2002) noted, this comparative fit index is independent of both sample size and model complexity and uncorrelated with overall fit measures.

By this standard, the metric and scalar invariance of the AAI was corroborated in the present study, and the factors of this scale were deemed to reflect equivalent domain representations in boys and girls from Grades 8 to 12. According to Meredith’s (1993) criteria, strong invariance was thus documented by the results of the present study. This was gratifying considering that it is increasingly argued that factorial invariance testing must include manifest means (Little, 1997).

By increasing our understanding of the psychometric properties of the AAI, the present study also expands our fundamental knowledge of the structure of academic amotivation. This concept has traditionally been conceived to be unidimensional (e.g., Deci & Ryan, 1985, 2002). Results from a prior study pertaining to ecological motivation suggested that amotivation to engage in proenvironmental behaviors adopted a multifactorial structure (Pelletier et al., 1999). Legault and her colleagues (2006) also demonstrated that academic amotivation could be advantageously conceptualized as a four-factor phenomenon. The present findings offer a replication and an extension of those results. That is, the four-factor amotivation taxonomy proposed by Legault and colleagues was found to be equally relevant and applicable for students of both genders for all high school grade levels.

In addition to the multigroup confirmatory factor analyses devised to test the equivalency of factor loadings and intercepts across grade and gender, invariance testing was also performed to examine differences in latent means. Gender differences were obtained for a single amotivation dimension: effort beliefs. Boys scored higher than girls in all grades, and the magnitude of their factor scores increased over grades. That is, in our sample, boys felt that they had less capacity or less willingness to invest effort in school tasks than did girls. These results are congruent with past findings of increased motivational deficits in boys (Grouzet et al., 2006; Ratelle et al., 1994; Vallerand et al., 1997).

It is intriguing that no other gender differences were found in the present study. Does this indicate that gender differences in amotivation are solely channeled through the aversion of scholastic exertion? This is a distinct possibility. It has been noted that, by comparison to girls, boys are more reluctant to do schoolwork, less determined to solve difficult problems, and less productive (Martin, 2004; Rowe & Rowe, 1999). As for amotivation due to poor capacity beliefs, insufficient academic values, and unappealing task characteristics, it is perhaps unsurprising that no gender differences were observed, considering the large amount of contemporary social emphasis on gender equality in school and career achievement and pursuits. In North
America nowadays, feeling competent in school and finding value and interest in education (or failing to do so) may be tied to factors that are independent of gender. It is therefore possible that our results reflect contemporary teenage attitudes accurately and that gender differences in motivation deregulation arise primarily from effort beliefs. Alternatively, our results may be sample specific. Motivational dynamics are grounded in a wide array of social and personal factors, and gender differences are intricate issues embedded within the complex process of human development. Substantial future research on this topic is needed before definite conclusions can be drawn. Future studies on this issue are of focal importance, as they may shed some light on the reasons why boys are more vulnerable to detrimental school outcomes, including dropout, than are girls (Janosz, 2000; Janosz et al., 1997; Martin, 2004).

As expected, it was possible to identify grade effects within amotivation dimensions. For boys and girls, amotivation due to insufficient school values and amotivation resulting from unappealing characteristics of school tasks increased over grade. The same pattern was observed in boys for amotivation due to lack of effort beliefs. These results are consistent with those of many studies that found that motivation decreases as students progress within the high school system (Eccles et al., 1993; Grouzet et al., 2006; Harter et al., 1992; Otis et al., 2005; Peetsma et al., 2005) and suggest that such decreases in motivation over time may represent manifold expressions of behavior regulation deficits. It is noteworthy that amotivation due to low capacity beliefs was stable across all grades. This could denote that ability beliefs are established early on and are stable over time. Long-term longitudinal studies would be required to fully test this hypothesis.

The magnitude of average factor scores was finally compared across amotivation dimensions. No differences were found in Grade 8. For all other grades, amotivation due to insufficient values (boys and girls) and amotivation due to a lack of effort beliefs (boys only) were higher than amotivation due to low capacity beliefs. Amotivation due to unappealing task characteristics displayed the highest latent means of all amotivation dimensions for both genders. Low perceived ability has long been considered to play a central role in negative school outcomes. Yet the present results connote that other cognitive and affective dimensions may be as important, perhaps even more so.

A better understanding of the dimensionality of amotivation may therefore be useful in clarifying the complex dynamic involved in the deregulation of school-related attitudes and behaviors, thereby facilitating the design and implementation of intervention programs for disengaged students. Our results suggest the need for widespread academic intervention and prevention designed to target the multifarious manifestations of school amotivation. Specifically, such programs could endeavor to increase students' competence and beliefs in their academic abilities, boost students' beliefs about their academic effort capacity, help students enhance the value and worth of school, and finally amplify the interesting and stimulating characteristics of academic activities and assignments. Our results also suggest that
it could be advantageous to intervene more actively in more advanced grades and to pay special attention to effort management in boys.

It must be acknowledged that the assessment of the internal consistency of AAI scores yielded a few values that were somewhat below the commonly accepted .80 criteria (Henson, 2001). These lower coefficients were all obtained for the effort belief subscale and may possibly be explained by the nature of this subscale. Effort is a construct that comprises an important behavioral element. By contrast, ability beliefs, school values, and task characteristics possess more pronounced cognitive and affective components. It is typically easier to achieve high levels of internal consistency for cognitive and affective constructs. Behavioral constructs are harder to assess reliably because behaviors are more idiosyncratic (i.e., person specific). However, regardless of the source of greater heterogeneity in this subscale, it would be useful to attempt to improve on its extant version and to evaluate it anew in future studies.

It must also be noted that the present study is limited by the cross-sectional nature of its design. It would be interesting to expand our knowledge of the functional equivalency of the AAI by testing its invariance across time in future longitudinal research projects comprising several data waves. A recent study successfully documented the factorial invariance of the Academic Motivation Scale (AMS; Valier-and et al., 1993) across gender and time (one wave per year during 3 years) in junior high school students (Grouzet et al., 2006). It would be interesting to ascertain whether the structure of amotivation is as stable over time as that of the more positive forms of behavior regulation that are assessed by the AMS.

References


