# Beyond HEQCO's Skills Assessments: Validation of a Measure of Students' Academic Capital and Evidence of Grade Inflation in Ontario's High Schools and Universities

James E. Côté[1], J. Paul Grayson[2], Sharon Roberts[3], & Liang Hsuan Chen[4]
[1]Western University, [2]York University, [3]University of Waterloo,
[4]University of Toronto

## Abstract

In the U.S., it is well documented that many students enter universities unprepared, lacking the basic academic skills necessary for optimal, or even positive, outcomes. However, less evidence has been reported on this problem in Canada, and there appears to be a public impression that Canada's universities have higher standards, on average, than American universities. This perception is reflected in documents produced by the Ontario Government, but documents prepared by independent sources, such as Statistics Canada and the Higher Education Quality Council of Ontario (HEQCO), suggest otherwise.

The present article adds to this independent literature, reporting on a new empirical measure based on a suite of questions assessing the basic academic skills necessary to meet the challenges of higher education. Seven sets of skills were identified, and through exploratory and confirmatory factor analyses, two forms of academic capital were identified, as were three clusters of students. Evidence for the validity of these measures was found in terms of hypothesized associations with student grades, satisfaction, and thoughts of dropping out. Demographic differences (e.g., family background, gender) were minimal. The results suggest that a majority of Ontario's university students are at risk of sub-optimal academic outcomes, including a sizeable minority that is likely dysfunctional in contexts in which using basic academic skills would be necessary to pass courses if it were not for grade inflation. These students report lower grades, greater dissatisfaction with the university experience, and more frequent thoughts of dropping out. Policy implications are discussed in terms of what governments, secondary schools, and universities might do to reduce skill deficiencies and the associated negative experiences with post-secondary education.

*Keywords:* university students, academic skills, skill deficiencies, grade inflation, student dissatisfaction, student dropout

## The General Policy Context: Academic Skills Development

From kindergarten to grade 12, one of the objectives of the Ontario Ministry of Education is "to support high-quality learning while giving individual students the opportunity to choose programs that suit their skills and interests" (Ministry of Education, 2018, p. 5). The Ministry makes it clear that it is important for students to develop various skills and work habits "*in preparation for postsecondary education* [our emphasis] and the world of work" (Ministry of Education, 2010, p. 12). These skills and work habits include: responsibility (e.g., takes responsibility for and manages own behaviour), organization (e.g., devises and follows a plan and process for completing work and tasks), independent work (e.g., follows instructions with minimal supervision), collaboration (e.g., accepts various roles and equitable share of work in a group), initiative (e.g., looks for and acts on new ideas and opportunities for learning), and

self-regulation (e.g., sets own individual goals and monitors progress towards achieving them) (Ministry of Education, 2010, p. 11). (Further elaboration, and operationalization of these skills, is found on pages 24 and 25 of Ministry of Education, 2010). These skills are consistent with the assessments of university instructors and employers regarding the factors that contribute to success in other jurisdictions (Arum & Roska, 2011; Tsui, 2002).

Despite these laudable objectives, in Canada in general, and in Ontario in particular, there is a belief among many university faculty members (Côté & Allahar, 2007, 2011; Dion & Maldonado, 2013; Hutchins, 2017), employers (Mancuso, 2014 ), Statistics Canada (Munro, 2014; Statistics Canada, 2016), think tanks (Mahboui, 2017), and journalists (Bradshaw, 2011; Brown, 2016) that many students entering universities lack the skills essential to academic success and future employment.

## HEQCO's Essential Adult Skills Initiative

Consistent with these concerns, in Ontario, steps have been taken to obtain a clearer picture of undergraduates' *literacy, numeracy,* and *problem-solving* skills by the Higher Education Quality Council of Ontario (HEQCO). In co-operation with several universities and a number of colleges, in 2013, HEQCO announced the intent of developing a cross-sectional study to measure university students' academic skills upon entry to, and exit from, university (Weingarten, 2014). By 2017, 11 colleges and 11 universities in Ontario were affiliated with the project (HEQCO, 2017). Preliminary results of this endeavour show that gains in these skills between university matriculation and graduation are small, and that one-quarter of graduates have low levels of these skills (Weingarten & Hicks, 2018). Although that report does not mention correlations of these skill measures with grades, it is possible that low-skilled students are graduating because of grade inflation associated with institutional pressures to graduate as many students as possible, a possibility suggested by the results of the present study discussed below.

In a second, longitudinal project designed to measure university and college students' *analytic and synthetic* skills, HEQCO conducted additional research in one college and one university in Ontario. Perhaps most important are their findings that there was only a very slight increase in *analytic and synthetic skills* among university students, and that scores of entering and graduating college students indicated no improvement (Finnie et al., 2018).

## Limitations of HEQCO's Skills Assessments

While studies such as those carried out by HEQCO are important, they have their limitations:

1. Given that public access to the questions on which the studies are based is limited, researchers cannot assess the overall adequacy of HEQCO's operationalizations of skills.
2. The operationalizations for the cross-sectional study to which researchers do have access seem remote from what is needed for students to negotiate academic careers, or what would be developed over the course of post-secondary studies.
3. The questions from the longitudinal study to which researchers have access suggest a possible disadvantage for students specializing in certain areas of study.
4. No Canadian studies have been undertaken that confirm a link between the skills measured in the two studies and academic success.
5. American evidence suggests that cognitively taxing questions, such as those involved in the two HEQCO studies, do not elicit respondents' best efforts (Liu et al., 2016).
6. There are some important abilities that simply cannot be measured readily through the vehicles employed by HEQCO. They include time management, carrying out research, debating and presentation skills, and group management.
7. These studies do not provide direct evidence regarding how grade inflation may be affecting the academic proficiencies of university students.

In essence, HEQCO identified deficiencies in certain skill sets of Ontario's undergraduates. However, the ways in which skills were operationalized is not clear, and the relationship of these skills to academic success is assumed, but not demonstrated.

## The Present Study

In view of the above concerns, the current study was driven by four research questions. First, can specific academic skills be reliably and validly measured? Second, what is the magnitude of deficiencies in these skills among Ontario's university students? Third, what are the policy implications of the relationships between skill deficiencies and measures of academic success? And fourth, what steps might be taken to help students overcome these skill deficiencies?

We proceeded with this investigation as follows. First, having recognized the limitations of the important work completed by HEQCO, and on the basis of our experience and reviews of the relevant literature, we identified skills that likely have a direct connection to academic success. These skills were operationalized in objective ways easily understood by undergraduates (i.e., specific, identifiable tasks involving the following generic academic skills: test-taking, conducting literature searches, analyzing researched material, writing essays, giving oral presentations, managing time and working with other students, and numeracy). We refer to these abilities as "academic capital," characterized as practices that become embodied in the person as mental capacities, through which the person can understand, research, abstract, critically analyze, create, and communicate various forms of knowledge. We then assessed the extent to which students possessed academic capital and the degree to which this form of capital is related to specific measurable outcomes: academic achievement (grades), retention (thoughts of leaving university before graduation), and satisfaction with the university experience.

In developing the skills items in terms of face validity, we adopted the approach frequently employed in occupational training. First, we identified the specific skills needed to meet the demands of the job (in this case, course work). Earlier referenced literature (Arum & Roska, 2011; Tsui, 2002) and our own experiences provided referents. Second, we measured the difference between what students could do and what they should be able to do to ensure academic success. We then determined, based on students' own assessments, if those possessing these skills were more academically successful than others. Finally, we designed curricular material to close the gap between what students should be able to do and their actual capabilities. In occupational settings, this approach is identified as "training to the gap." In evaluations, this approach includes observation of actual desired behaviours as well as traditional measures, such as performance on tests.

In addition to the face validity of the individual items, we evaluated the skills items in terms of two forms of validity: factorial and predictive criterion validity (Anastasi, 1988). As will be seen below, evidence for factorial validity is provided by the results of confirmatory factor analysis using structural equation modeling regarding the predicted subscales of the skills items. Evidence for the predictive criterion validity of the skills measures is found to the extent that the summative scores of the skills subscales are shown to be associated with higher student grades, more satisfaction, and fewer thoughts of dropping out.

In our analysis of skills in relation to grades, we measured both high school grades and university grades. As discussed below, we found an unexpectedly large discrepancy between these two levels of academic achievement, with high school grades significantly higher than university grades for the overall sample. This difference also varied by university. Collins (1998) referred to this discrepancy as "grade drop." He reported that in the mid 1990s the grade drop at Western was 14%, from 79.5% at the end of high school to 65.3% at the end of the first year. Although an additional study does not reference change over time, information collected at the University of Waterloo is consistent with this finding (Cain, 2018). From 2016 to 2018 the engineering faculty at Waterloo kept tabs on the first-year performance of students from 73 Ontario high schools. On average, first-year grades were 16% lower than high school marks (with a range of 10% to 28%). This difference was called the "adjustment factor."

Although the literature on grade drop is not extensive, given evidence of discrepancies between the two grade attainments in our samples, we calculated the variable "grade drop" by subtracting university grades from high school grades. Whereas we assumed lower grades at both high school and university would be related to lower skill levels, we also assumed that the greater the discrepancy between the two, the less prepared the student would be for university studies in terms of academic capital. That is, students with more extreme grade drops likely did not develop academic skills at high school commensurate with the grades they received. Moreover, as will be seen in the presentation of the results, we interpreted the unexpected extent of the grade drop as evidence of high school grade inflation.

Retention is an important issue in a number of respects. Students who leave a given university prior to degree completion represent a financial loss to that university (e.g., recruitment costs). As a result, it is in the interests of universities to retain students. For the student, a premature departure from school may have several consequences, including a lost opportunity for the development of human capital, a loss of future earnings, and costs already incurred to cover educational expenses. Accordingly, it is important to assess the extent to which student skills deficits are related to thoughts of leaving the university before graduating.

Similarly, student satisfaction with the university experience is a desirable outcome in its own right. Satisfaction has also been associated with a number of positive university outcomes, such as alumni financial support. Moreover, various agencies collect and publish the degree of satisfaction reported by undergraduates at various universities. These measures are often used to rank, and thereafter market, particular institutions to potential students. Consequently, from an institutional perspective, it is important to maximize students' satisfaction with the university experience. Accordingly, it is important to assess the extent to which student skills deficits are related to levels of satisfaction.

## The Sample and Procedure

In total, 1,995 students from humanities, social sciences, and some professional studies faculties at York University, the University of Waterloo, Western University, and the University of Toronto (Scarborough and Mississauga campuses only, which were treated as one group) responded to an online survey asking them to rate their competence and confidence in key academic skills. All four samples were surveyed in the mid-fall term. The sample sizes (*n*s) for each institution were 949, 278, 203, and 565, respectively.

In general, missing data resulted from respondents discontinuing the survey. Given the large size of this sample, these cases were deleted listwise for the particular analysis at hand. Missing data for the 48 skills questions were found for only 12 cases for the first question but increased toward the latter questions. The lowest *n*s were found in the multivariate analyses (dropping to 1467) that included demographical variables (it was not feasible to replace missing data on variables such as gender, international student status, and so forth).

Demographic characteristics of the final, overall sample were as follows (with 9% missing overall): 12% of the sample were international students, 40% were classified as first-generation university attenders (based on reported parents' education, these students were the first in their family to attend university), 28% spoke a language other than English as a child, and 64% identified as female (24% males, 3% "other").

### *Analysis Plan*

The major concept of interest, academic capital, was first assessed in terms of factorial validity by treating the 48 academic skills items as seven parcels, namely, the sum of each of seven proposed subscales (Byrne, 2010). Then, the seven subscales were subjected to cluster analysis to explore possible typologies representing patterns in academic capital attainment.

Next, the predictive validity of the factors and clusters were assessed in terms of their relationship with high school and university grades (as well as grade drop), along with two negative student experiences that university administrators have been at pains to reduce: (1) dissatisfaction with the university experience, and (2) thoughts of dropping out. These assessments were undertaken with a series of bivariate and multivariate analyses.

In evaluating effect sizes, following Cohen (1988), correlations less than .10 are considered trivial, those between .10 and .30 are considered small effects, those between .31 and .50 are considered moderate effects, and those greater than .50 are considered large effects. Similarly, variance explained ($R^2$ and $\eta^2$) less that 1% (.01) is considered trivial, less than 10% (.10) small, less than 25% moderate, and greater than that large. In terms of comparisons between two means, Cohen's *d*s less than .20 are considered trivial.

## Measures

### *Academic capital*

**Seven subscales.** As a first step, we identified seven forms of academic capital potentially contributing to academic success. Several examples of the ways in which these items might be operationalized were suggested in the academic literature (Williams & Stockade, 2003). We then operationalized specific aspects of academic capital to reflect the key skills important in mastering academic material in the humanities, social sciences, and some professional studies. In the questions we developed, students were asked to assess their ability and confidence to conduct literature searches (research), to analyze written material (analysis), to write coherent and grammatically correct papers (writing), to study for and pass different types of tests (test-taking), to prepare for and give presentations (presentations), to carry out elementary numeric calculations (numeracy), and to manage their time and their learning groups (time & group management).

The 48 questions making up the seven subscales of academic capital constituted the core of the survey. All questions focused on students' self-assessments of their ability to utilize specific forms of academic capital in their courses. All questions had a fixed 5-point Likert format, on either an easy–difficult or a confident–unconfident range, with higher scores indicating that respondents found the exercise of each skill more difficult or lacked confidence in performing a skill-based task. In other words, higher scores on each item reflected a specific deficit in academic capital.

### *Academic Capital: Exploratory and Confirmatory Factor Analysis*

Each of the seven groups of items identified above was factor analyzed separately using the maximum likelihood estimate, and each proved to constitute one factor with acceptable Cronbach's alpha reliabilities, ranging from .75 to .89. Scores for each of the seven subscales were calculated by computing the "mean of mean" so that each subscale had equal ranges (and therefore weights) of 1 to 5.

The seven subscales were then subjected to an exploratory factor analysis (EFA) and confirmatory factor analysis (CFA).

In the EFA, using only the respondents from York University (where the survey was first administered in the fall of 2017; $n = 855$), the seven summed skills subscales were treated as "parcels" (Byrne, 2010) using maximum likelihood extraction and varimax rotation. Only one factor was extracted explaining 54.88% of the variance (Eigenvalue = 3.84). The Kaiser-Meyer-Olkin Measure of Sampling Adequacy was .91 ($p < .000$). The factor loadings for each subscales were as follows: research = .81, analysis = .85, writing = .90, test taking = .75, presentations = .70, numeracy = .40, and time & group management = .65.

Then, using the respondents from the other three universities (surveyed in the fall of 2018; $n = 832$), a CFA for a one-factor model was performed with Amos version 25 using the maximum likelihood estimator. The coefficients for the steps taken to improve the measurement model are presented in Table 1. The modification indices (MIs) in the first step indicated problems with the (low) regression weight for the numeracy subscale. After dropping that subscale, the model improved, but the MIs suggested that a covariance should be added for the error terms between research and test-taking. The model improved with this covariance, but the MIs indicated another covariance should be added for the error terms between presentations and time & group management. After adding this covariance, a good fit was found (CMIN/DF < 5.0; RMSEA < .05; CFI > .95), and the MIs did not suggest that any further improvement could be made to the measurement model.

Two concerns remained, however. First, the PCLOSE was slightly below the .50 cut-off signifying the RMSEA is actually < .05. Second, three of the regression weights/factor loadings were less than .70. Accordingly, a two-factor measurement model was tested, with the three highest-loading subscales designated to correspond to a latent factor "abstract academic capital" and the three lowest-loading subscales (along with numeracy) designated to correspond to a latent factor "concrete academic capital."

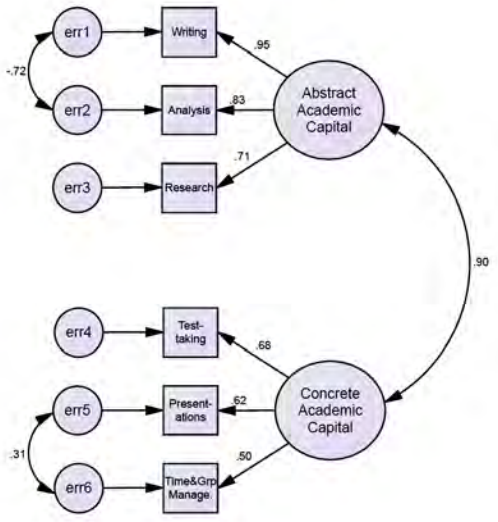
**Table 1**
*CFA Modification Indices and Steps in Model Fitting for the Latent Factor Academic Capital*

| Model | CMIN/DF | RMSEA | ΔRMSEA | PCLOSE | CFI | ΔCFI |
|---|---|---|---|---|---|---|
| 1-Factor | 11.283 | .108 | -- | .000 | .936 | -- |
| Modification 1 – drop numeracy | 15.681 | .129 | .021 | .000 | .940 | .004 |
| Modification 2 – cov. err5 & err6 | 5.164 | .069 | .060 | .059 | .985 | .045 |
| Modification 3 – cov. err3 & err4 | 3.085 | .049 | .020 | .493 | .993 | .008 |
| 2-Factor | 10.015 | .093 | -- | .000 | .956 | -- |
| Modification 1 – cov. err5 & err6 | 5.744 | .067 | .026 | .030 | .978 | .022 |
| Modification 2 – cov. err6 & err7 | 5.432 | .046 | .021 | .671 | .982 | .004 |
| Modification 3 – drop numeracy | 5.793 | .050 | .004 | .473 | .994 | .012 |
| Modification 4 – cov. err1 & err2 | 2.39 | .041 | .009 | .670 | .996 | .002 |

Table 1 also shows the steps in developing a two-factor measurement model, which had a slightly better fit than the one-factor solution. MIs first suggested adding covariances between the error terms for presentations and time & group management, and then time & group management and numeracy. However, when this was done, the MIs suggested that numeracy should be dropped. When these changes were implemented, the model improved, but the MIs suggested adding a covariance between the error terms for writing and analysis (err1 and err2). This model showed a very good fit (CMIN/DF < 3.0; RMSEA < .05; PCLOSE > .50; CFI > .95) that would not be improved by any further modifications.

Although good fits can be found for both the one- and two-factor solutions, the two-factor model allows for the analysis of two manifestations of academic capital, and it identifies a shorter version of an academic capital scale for studies in which fewer questions can be asked for practical reasons (i.e., the short version can constitute just the three subscales with strong loadings comprising abstract academic capital). Figure 1 shows this solution, where the three highest-loading subscales correspond to a latent factor "abstract academic capital," and the three lowest-loading subscales correspond to a latent factor "concrete academic capital."

**Figure 1**
*The CFA Results Showing the Two-factor Model, Along with Standardized Regression Weights and Error Covariances.*



**Cluster analysis.** Whereas factor analysis provides information about how variables group together (a variable-centered analysis), cluster analysis provides information about how people group together (a person-centered analysis). To examine how academic capital varies among groups of respondents, the seven subscales were subjected to a *two-step cluster analysis.* This procedure takes the information from all variables (mean scores summed in the seven scales) and enters them into an algorithm to produce a typology based on how those variables cluster with each other to form groups of like-scoring respondents. We call the results of this procedure the *skill-proficiency typology.* With the data from the respondents from all universities combined, this procedure produced a three-group typology: 44% were clustered into a "functional" group, 41% were classified as "at-risk," and the remaining 15% were categorized as "dysfunctional." The quality analysis of this solution has a fair "silhouette measure of cohesion and separation," with mean differences among the groups having large effect sizes on all of the individual skills subscales ($\eta^2$ = .37 to .61), except the numeracy subscale, which had a moderate effect size ($\eta^2$ = .12).

### *Grades*

Given the common assumption that grades reflect student learning and skill development, we measured high school grades and current university grades on 8-point scales ranging from D to A+. Respondents were asked to estimate their overall grade in their final year of high school, as well as what they thought their overall average would be during their current academic year at university. Although it would have been preferable to use administrative data as an estimate of student grades, research based on multiple samples has found self-report and administrative recorded grades to be highly correlated ($r$ = .90 for university grades and .82 for high school grades; Kuncel et al., 2005). Because these are single-item measures, it is not possible to calculate Cronbach's alpha reliability coefficients for either grade estimate.

Given the evidence of discrepancies between the two grade attainments for many students (e.g., Collins, 1998), we calculated the variable "grade drop" by subtracting university grades from high school grades. This variable is an estimate of different grading standards between high school and university, with university grades providing the reference point for academic standards. Accordingly, large positive differences between these two grades were interpreted as evidence of students receiving inflated grades at high school.

### *Thoughts of Dropping Out*

Subjects were asked, "Some students get fed up and leave the university before they complete their degrees. Others stay until degree completion. How often, if ever, do you feel fed up and think you would just like to quit?" They were provided a four-point scale to answer ("never," "sometimes," "often," "all the time"), such that higher scores reflect more frequent thoughts of dropping out.

### *Dissatisfaction with the University Experience*

Subjects were asked, "Overall, how satisfied are you with your experience at [your current university] so far?" and were given a five-point scale to answer ("very satisfied" to "very dissatisfied"). Thus, higher scores reflect greater dissatisfaction.

## Results

In order to estimate the extent of skills deficiencies for each of the 48 items, prior to the factor and cluster analyses, each was recoded into a "0/1" dummy variable format, with "1" representing the student reporting a difficulty or lack of confidence with the skill in question (i.e., "4" or "5" on the five-point scale were recoded to "1"). Those who selected "3" (neutral) were given the benefit of the doubt for being proficient. As a result, our findings are likely conservative estimates of risk and dysfunction. Consistent with the foregoing coding practices, Appendix A contains the 48 items, with the first column identifying the skill question under consideration, and the remaining columns showing the percentage of students at each institution expressing difficulty or lack of confidence in their mastery of the skill.

Overall, percentages in Appendix A point to a great deal of uniformity among the universities. For example, 26% of students at Waterloo reported difficulty with writing a one-page summary of an article. The highest figure for this item (35%) was recorded for Toronto. Overall, 30% of students reported dif-

ficulty with the task.

When examining the percentages in Appendix A, it can be seen that students reported the greatest difficulty in identifying the types of arguments embodied in articles and books (analysis subscale). Despite minor fluctuations from one university to the next, overall, 57% of students reported that for them the task was difficult or very difficult. The second most frequently cited deficiency was writing a critical book review (writing subscale). Forty-nine percent (49%) of students recognized difficulties with this task. In contrast, the least difficult task was simple subtraction, with only a small percentage (8%) rating it as difficult (numeracy subscale).

## *Bivariate Analyses*

**Skill-proficiency typology.** Table 2 shows the relationships between the skill-proficiency typology generated via the cluster analysis and the various measures. As one would expect, the three clusters clearly differentiate the two forms of academic capital, with very large effects sizes. In terms of mean scores on the 1 to 5 scales, the functional group rated their skills at a mean of approximately "2" corresponding to easy/confident; the at-risk group rated their skills at a mean that rounds to "3" (neutral); and the dysfunctional group rated their skills at a mean that rounds to "4" (difficult/unconfident).

**Table 2**
*Skill-proficiency Typology Cluster Differences on Key Measures: One-way ANOVAs, Means, Sds, Multiple Range Tests, and Effect Sizes*

| | Skill-Proficiency Typology | | | | |
|---|---|---|---|---|---|
| Measure | Functional (*n*s=745-796) | At-risk (*n*s=699-742) | Dysfunctional (*n*s=263-282) | *F*-value | $\eta^2$ |
| Abstract Academic Capital Deficiency | $2.29_a$ (0.45) | $2.84_b$ (0.41) | $3.62_c$ (0.54) | 956.23*** | .60 |
| Concrete Academic Capital Deficiency | $2.27_a$ (0.48) | $2.86_b$ (0.44) | $3.67_c$ (0.57) | 947.55*** | .51 |
| High School Grades | $6.59_a$ (1.78) | $6.46_a$ (1.15) | $6.16_b$ (1.29) | 13.34*** | .02 |
| University Grades | $5.89_a$ (1.21) | $5.39_b$ (1.37) | $4.65_c$ (1.45) | 94.85*** | .10 |
| Grade drop | $0.73_a$ (1.40) | $1.09_b$ (1.66) | $1.54_c$ (1.74) | 28.25*** | .03 |
| Thoughts of dropping out | $1.66_a$ (0.81) | $2.06_b$ (0.97) | $2.46_c$ (1.08) | 88.87*** | .09 |
| Dissatisfaction | $2.22_a$ (0.91) | $2.64_b$ (1.00) | $3.08_c$ (1.11) | 86.92*** | .09 |

*Note.* Higher scores on the two capitals are indicative of greater deficiency. Means that do not share subscripts differ by $p < .05$ according to the Student-Newman-Keuls multiple-range tests. Missing cases deleted case-wise, so *n*s reported above reflect the range of cases for each of the seven variables reported for each skill-proficiency cluster. * $p < .05$; ** $p < .01$. *** $p < .001$.

The three skills-proficiency groups are clearly differentiated on all of the other measures. The dysfunctional group had the most problematic profile: the lowest grades, greatest grade drop, more frequent thoughts of dropping out, and most dissatisfaction. In contrast, the functional group had the most favourable profile, with the at-risk group scoring in between both groups on all measures.

With respect to reported grades, it is worth noting that the grade level "6" corresponds to a B+, so it is striking that all three skill-proficiency groups graduated high school within the B+ range in spite of their significantly different skill levels. The grade drop level of 1.5 for the dysfunctional group corresponds to a decrease of three-quarters of a full letter grade (C+), suggesting that these students likely received inflated high school grades. At the same time, the finding that those in the dysfunctional group were receiving a C+ average at university suggests that they were also receiving inflated university grades. Grade inflation would explain why students with very low skill levels would be awarded a grade (C) that was traditionally the average grade awarded in universities and which signified a "satisfactory" performance (Anglin & Meng, 2000; Levine & Curetan, 1998).

In other words, those in the dysfunctional group rate themselves as having low levels of the most basic academic skills, so they presumably lack the competence to pass tests and assignments in which

traditional standards are used, yet they are being awarded university grades signifying that their performances are satisfactory. Grade inflation in high school is likely one of several means of encouraging more students to stick it out to earn a diploma in order to qualify for more jobs (e.g., Côté & Allahar, 2011); similarly, grade inflation at university is likely one of several means to increase retention (Anglin & Meng, 2000), which to some extent is in the interests of both students (earning a credential that differentiates them from high school graduates) and universities (in terms of tuition revenues, given current government funding levels).

**Inter-university comparisons.** The four universities showed statistically significant differences on many of the measures, although some of these differences were minor, with small effect sizes (and were likely significant because of the large sample size, as in the case of the two forms of academic capital, thoughts of dropping out, and dissatisfaction, where the F-values are relatively low). Table 3 shows the results of the one-way ANOVAs and the SNK multiple-range tests for these comparisons (no differences were found on any measures between the two Toronto campuses, so they were combined in the analyses).

**Table 3**

*Inter-University Differences on Key Measures: One-Way ANOVAs, Means, Sds, Multiple Range Tests, and Effect Sizes*

| | University | | | | | |
|---|---|---|---|---|---|---|
| Measure | York (*n*s=866-947) | Toronto (*n*s=538-562) | Waterloo (*n*s=174-278) | Western (*n*s=135-203) | *F*-value | $\eta^2$ |
| Abstract Academic Capital Deficiency | $2.76_a$ (.68) | $2.75_a$ (.63) | $2.63_b$ (.60) | $2.68_a$ (.68) | 3.19* | .01 |
| Concrete Academic Capital Deficiency | $2.74_a$ (.71) | $2.76_a$ (.67) | $2.59_b$ (.62) | $2.75_a$ (.68) | 3.38* | .01 |
| High School Grades | $6.14_a$ (1.25) | $6.70_b$ (1.09) | $6.94_c$ (0.80) | $7.20_d$ (0.89) | 60.46*** | .09 |
| University Grades | $5.56_a$ (1.30) | $5.20_b$ (1.54) | $5.84_c$ (1.14) | $5.66_{ac}$ (1.36) | 14.09*** | .02 |
| Grade drop | $0.60_a$ (1.53) | $1.51_b$ (1.69) | $1.09_c$ (1.20) | $1.53_b$ (1.26) | 44.70*** | .07 |
| Thoughts of dropping out | $1.89_a$ (0.95) | $2.08_a$ (1.0) | $1.96_a$ (0.94) | $1.89_a$ (0.88) | 4.57** | .01 |
| Dissatisfaction | $2.52_a$ (1.05) | $2.63_{ab}$ (1.05) | $2.44_{ab}$ (0.98) | $2.34_b$ (0.92) | 3.81** | .01 |

*Note.* Means that do not share subscripts differ by $p < .05$ according to the Student-Newman-Keuls multiple-range tests. Missing cases deleted case-wise, so *n*s reported above reflect the range of cases for each of the seven variables reported for each school. * $p < .05$; ** $p < .01$. *** $p < .001$.

The most notable differences were for grades. York students reported the lowest high school grades, whereas Western students reported the highest. These results are consistent with institutional data, with York's reported incoming student average reflecting the lowest entrance standard of the four schools (82%), Western and Waterloo the highest (90%), and Toronto in-between (87%) (these are average final-year grades of full-time, first-year students entering university from high school; Dwyer, 2017). Given the overall pattern of results of this study, it is possible that many students in the three more "selective" universities actually received the most inflated high school grades (for reasons we cannot assess with our data). The extent and differences in grade inflation can be seen in grade drop from high school to university: Western and Toronto students reported the greatest drop—1.53 points on the scale—and York students the lowest drop, suggesting that students in the former schools received more inflated high school grades than students in the latter school or that there are some mechanisms in place (higher

standards) to attempt to curb grade inflation.

University grades showed a different pattern, with Toronto students reporting the lowest earned grades and Waterloo students the highest, with Western and York in between.

The grade drop results from Western are particularly interesting, given Collins (1998) documentation of a 14% grade drop in the mid 1990s at Western. At that time, the incoming average was 79.5% (a B+), and first-year final grades averaged to 64.5%. However, 20 years later, our results suggest that Western's incoming average is well into the As (7.20 on our scale). The entrance grade average at Western is now topping the 90% level, with 55% of incoming students having A+ high school graduation averages according to online administrative data: http://www.ipb.uwo.ca/pi/secondary_school_grades_of_incoming_students.php (42.2% of first-year Western students in our sample reported an A+ high school average). When just first-year students are examined in our Western sample, the university average is now a solid B (5.26), so the grade drop is about the same, but from two higher levels (from A to B), suggesting significant grade inflation at the university level as well as the high school level over the past 20 years.

Although not presented in a table, the four universities were also compared in terms of their respective proportions of the three skill-proficiency groups. Using cross-tabulation analysis, two universities stood out in terms of observed/expected frequencies, although the overall effect size was trivial ($\eta 2 < .01$). York had significantly more functional students than expected ($z = 2.6$, $p < .05$), and fewer at-risk students ($z = -3.4$, $p < .05$); Waterloo had more at-risk students than expected ($z = 2.3$, $p < .05$), but fewer dysfunctional students ($z = -3.4$, $p < .05$). No university stood out as having more students with extreme academic capital deficits.

**Year level.** One-way ANOVAs were also performed on year level. It can be seen in Table 4 that academic capital deficiencies decline slightly with year level, more so for abstract than concrete capital (consistent with HEQCO results reported by Finnie et al., 2018). The multiple-range tests indicate that first-year students report the greatest deficiency and that the level of deficiency decreases with a small effect size with each year level, such that those in the two highest years (fourth and fourth+) report the lowest deficiencies. With respect to concrete academic capital, although the F-test is significant, the multiple-range test does not differentiate among the year levels, and the effect size was trivial in terms of differences among the means (the fact that the F-test was significant is likely due to the large sample size). These results also need to be tempered with the separate cross-tabulation finding that there are no changes in the percentages in the skill proficiency groups over the five year-levels. This latter finding supports HEQCO's findings of no skill development reported by Weingarten and Hicks (2018).

**Table 4**

*Year-level Differences on Key Measures: One-way ANOVAs, Means, Sds, Multiple Range Tests, and Effect Sizes*

| | Year Level | | | | | | |
|---|---|---|---|---|---|---|---|
| Measure | First (*n*s=519-544) | Second (*n*s=410-433) | Third (*n*s=337-359) | Fourth (*n*s=195-237) | Fourth+ (*n*s=195-215) | *F*-value | η2 |
| Abstract Academic Capital Deficiency | $2.83_{a}$ (.62) | $2.77_{ab}$ (.58) | $2.69_{bc}$ (.73) | $2.56_{cd}$ (.68) | $2.60_{d}$ (.62) | 10.7*** | .02 |
| Concrete Academic Capital Deficiency | 2.78 (.65) | 2.76 (.62) | 2.72 (.75) | 2.63 (.71) | 2.67 (.74) | 2.36* | .00 |
| High School Grades | 6.53 (1.21) | 6.43 (1.17) | 6.44 (1.18) | 6.61 (1.11) | 6.43 (1.23) | 1.29 | .00 |
| University Grades | $5.20_{a}$ (1.46) | $5.49_{b}$ (1.22) | $5.66_{bc}$ (1.39) | $5.81_{ac}$ (1.35) | $5.51_{b}$ (1.36) | | .02 |

| | Year Level | | | | | | |
|---|---|---|---|---|---|---|---|
| Measure | First (*n*s=519-544) | Second (*n*s=410-433) | Third (*n*s=337-359) | Fourth (*n*s=195-237) | Fourth+ (*n*s=195-215) | *F*-value | η2 |
| Grade drop | $1.35_{a}$ (1.67) | $0.94_{b}$ (1.46) | $0.81_{b}$ (1.60) | $0.84_{b}$ (1.47) | $0.93_{b}$ (1.63) | 8.43*** | .02 |
| Thoughts of dropping out | $1.83_{a}$ (0.92) | $1.90_{ab}$ (0.92) | $2.01_{bc}$ (1.04) | $2.04_{bc}$ (0.93) | $2.19_{c}$ (1.06) | 6.63*** | .02 |
| Dissatisfaction | $2.41_{a}$ (0.95) | $2.57_{ab}$ (1.00) | $2.57_{ab}$ (1.09) | $2.50_{ab}$ (1.00) | $2.70_{b}$ (1.17) | 3.65** | .01 |

*Note*. Means that do not share subscripts differ by p < .05 according to the Student-Newman-Keuls multiple-range tests. Missing cases deleted case-wise, so ns reported above reflect the range of cases for each of the seven variables reported for each year level. * $p < .05$; ** $p < .01$. *** $p < .001$.

Table 4 also shows that thoughts of dropping out increase with year level. Those in the fourth+ year group score highest, whereas those in first-year score lowest. Dissatisfaction is also slightly higher among the fourth+ group. University grades also increase with year level, with first-year students reporting the lowest (at the B level) and fourth-year students the highest (still at the B level, but rounding up to a B+ level). Grade drop is greatest for those in first year, with the upper years not differing from each other, perhaps because these universities grant higher grades in the upper years (when university grades were controlled in an analysis of covariance, year level did not predict grade drop).

**Gender.** Differences emerged on all measures for gender, albeit small ones. As noted above, 96.7% of those answering the question identified as either male or female, with the remaining 3.3% identifying as LGBT or "other." This imbalance in the group sizes creates problems for multivariate analyses, with the 3.3% LGBT/other being dropped from the analysis. However, at the bivariate level, the results are more meaningful, although they need to be viewed with caution. Based on one-way ANOVAs and multiple range tests, compared with the male and female groups, those in the LBGT/other group reported a greater deficiency in concrete (but not abstract) academic capital, slightly higher thoughts of dropping out, greater grade drop, but higher high school grades. When just males and females are compared in t-tests, females reported greater deficiencies in concrete and abstract academic capital, slightly higher thoughts of dropping out, greater grade drop, and higher high school grades. However, all of the mean differences were trivial (i.e., all Cohen's *d*s < .20) and likely reached the level of significance due to the large sample size.

**First-generation university attenders.** A series of t-tests revealed no differences in the multiple-item measures between students with parents who had attended university and those without. The only differences were found for high school and university grades, with first-generation students reporting *slightly* lower grades in both cases.

**International students.** The profile of international students was relatively complex. For example, a series of t-tests revealed that they reported slightly lower high school grades than domestic students (although the effect size is trivial, Cohen's $d < .20$, so they can be included with the domestic students in analyses involving grades), but had slightly higher levels of abstract and concrete capital deficits. They also had fewer levels of thoughts of dropping out. (Note that these bivariate effects change significantly when the international student variable is assessed in regression analysis).

**English as a first language.** Although only one-third of those raised as children speaking a language other than English were international students, the profiles of the two groups were similar: slightly lower high school and university grades and higher levels of abstract and concrete capital deficits, but also lower levels of thoughts of dropping out.

**Correlations among measures.** Table 5 shows the correlations among the measures. The two capital deficit scales are associated with greater thoughts of dropping out and dissatisfaction, lower high school and university grades, and a greater grade drop from high school to university, all with small to moderate effects sizes. Thoughts of dropping out and dissatisfaction were unrelated to high school grades but were related to lower university grades and a greater grade drop.

The correlation pattern of high school grades with the other measures suggests that they are relatively poor predictors of all the other measures except grade drop. That is, high school grades do not predict the two negative student experiences (dissatisfaction and thoughts of dropping out) at all, and academic capital only a little. They do predict university grades somewhat better, but with a small effect size. However, they are a strong predictor of grade drop: the higher the grades assigned by high schools, the greater the drop at university ($r$ = .54); conversely, the lower the university grades, the greater the drop from high school ($r$ = -.69). One interpretation of this finding is that students with the highest high school grades were more likely to experience a larger grade drop at university. This may be due in part to a regression to the mean, but also because inflated high school grades do not prepare the students assigned those grades for the more rigorous grading standards at the university level (see Laurie, 2007, for evidence that inflated teacher assigned grades are associated with lower scores on objectively graded Provincial exams).

**Table 5**
*Pearson Correlations among the Measurement Scales*

| | Concrete Academic Capital Deficiency | High School Grades | University Grades | Grade drop | Thoughts of dropping out | Dis-satisfaction |
|---|---|---|---|---|---|---|
| Abstract Academic Capital Deficiency | .71*** | -.17*** | -.36*** | .20*** | .26*** | .30*** |
| Concrete Academic Capital Deficiency | -- | -.17*** | -.39*** | .22*** | .30*** | .35*** |
| High School Grades | | -- | .24*** | .54*** | -.03 | -.03 |
| University Grades | | | -- | -.69*** | -.32*** | -.32*** |
| Grade drop | | | | -- | .26*** | .25*** |
| Thoughts of dropping out | | | | | -- | .49*** |

*Note.* * $p < .05$; ** $p < .01$. *** $p < .001$.

**Regression analyses.** In two separate analyses, with thoughts of dropping out and dissatisfaction as dependent variables, the independent variables were entered in the following five blocks: (1) demographic, (2) university, (3) year level, (4) grades, and (5) academic capital deficiencies. Each was then repeated with the three skill-proficiency groups substituted in block 5 for academic capital deficiency measures in order to gain a sense of which form of reporting academic capital skills is more informative. In the analyses of the skill-proficiency groups, the functional cluster was the reference category, with the at-risk cluster and dysfunctional cluster each entered as dummy variables (0,1).

In block 1, demographic variables comprised: gender (0 = male; 1 = female), first-generation status (0

= second generation; 1 = first generation), international student status (0 = no; 1 = yes), and English as a first language (0 = English; 1 = non-English). Creating four separate variables for block 2, each university was dummy coded as 1, with the other three universities coded as 0. York was treated as the reference category, so a variable was not entered for it. In block 3, year at university was similarly coded into five variables representing each level. First year was treated as the reference category, so a variable was not entered for it in the equation. For block 4, although both high school and university grades were used, grade drop was not used because of multicollinearity problems with the two grade variables upon which it is based. Collinearity tests were performed on the other variables and no problems were found (i.e., none had a tolerance < .20, or a VIF > 4.0). Block 5 entered the academic capital measures as described in the preceding paragraph.

Table 6 shows the estimates of the final model with all five blocks entered in the order described above, regressed on *thoughts of dropping out*. In this model, all variables control for each other. It can be seen that many of the minor bivariate differences described above were washed out in this multivariate analysis. Of the demographic variables, only international student status and speaking a non-English language as a child remained significant controlling for the other variables. However, it can be seen in the block statistics that these demographic variables explained little variance ($R^2$ = .026). The next two blocks of variables also did not explain much variance: Toronto students had considered dropping out slightly more often than York students, and each year level represented slightly increased thoughts of dropping out over first year ($R^2$ of .008 and .010, respectively).

Most of the variance in the final model was explained by grades and the measures of academic capital deficiencies. Grades explained the most variance ($R^2$ change = .101), followed by the two academic capital deficiency scales ($R^2$ = .059). Looking at the estimates in Table 6, it can be seen that university grades had the largest beta (-.235), followed by abstract (.159) and then concrete (.134) academic capital deficiency. This said, with a combined beta .293, the effects of the capital variables actually exceed that of grades (-.235). Unexpectedly, high school grades were *positively* associated with thoughts of dropping out (i.e., the higher their grades coming out of high school, the more likely students were to consider dropping out of university, pointing to the possible negative effects of high school grade inflation). The negative relationship of university grades with thoughts of dropping out was as expected.

**Table 6**
*Regression Analysis on Thoughts of Dropping Out*

| | Estimates | | | | Blocks | | |
|---|---|---|---|---|---|---|---|
| Model 5 | *b* | *Std. Error* | *Beta* | *t* | $R^2$ | $\Delta R^2$ | $F\,\Delta$ |
| (Constant) | .957 | .277 | | 3.451** | | | |
| Block 1: Demographics | | | | | .026 | .026 | 10.01*** |
| Gender | .062 | .052 | .029 | 1.201 | | | |
| First Generation | -.007 | .046 | -.003 | -0.146 | | | |
| International student | .248 | .071 | .090 | 3.480** | | | |
| Childhood language | -.237 | .054 | -.114 | -4.402*** | | | |
| Block 2: Universities | | | | | .035 | .008 | 4.30** |
| Waterloo | .052 | .079 | .017 | 0.657 | | | |
| Western | .030 | .090 | .009 | 0.336 | | | |
| Toronto | .122 | .054 | .059 | 2.236* | | | |

| | Estimates | | | | Blocks | | |
|---|---|---|---|---|---|---|---|
| Model 5 | *b* | *Std. Error* | *Beta* | *t* | $R^2$ | $\Delta R^2$ | $F\,\Delta$ |
| Block 3: Year level | | | | | .045 | .010 | 3.97** |
| Year 2 | .130 | .062 | .057 | 2.083* | | | |
| Year 3 | .259 | .066 | .107 | 3.910*** | | | |
| Year 4 | .360 | .075 | .127 | 4.784*** | | | |
| Year 4+ | .420 | .079 | .141 | 5.339*** | | | |
| Block 4: Grades | | | | | .146 | .101 | 86.57*** |
| High school | .049 | .021 | .060 | 2.349* | | | |
| University | -.165 | .019 | -.235 | -8.810*** | | | |
| Block 5: Academic capital | | | | | .205 | .059 | 54.00*** |
| Abstract deficiency | .238 | .051 | .159 | 4.659*** | | | |
| Concrete deficiency | .190 | .049 | .134 | 3.892*** | | | |

*Note*. Listwise deletion, $N = 1467$; * $p < .05$; ** $p < .01$. *** $p < .001$.

Although not presented in table form, when the skill-proficiency clusters were substituted for the two academic capital scales in a regression analysis with the same independent variables entered in the same order, roughly the same amount of variance was explained by the overall model ($R^2 = .185$), and the clusters uniquely explained about the same amount of variance as the two scales ($R^2 = .051$). The beta for the at-risk group was .154, and for the dysfunctional group it was .241. Thus, either estimate of academic capital deficiency can be used in future studies, depending on the object of investigation.

Table 7 shows the equivalent regression results for dissatisfaction. The pattern generally follows that of the variable thoughts of dropping out, with many of the minor bivariate differences also washed out in this multivariate analysis. However, it can be seen in the *F*-values for the block statistics that neither the demographic nor the university variables were collectively significant in explaining variance.

**Table 7**
*Regression Analysis on Dissatisfaction*

| | Estimates | | | | Blocks | | |
|---|---|---|---|---|---|---|---|
| Model 5 | *b* | *Std Error* | *Beta* | *t* | $R^2$ | $\Delta R^2$ | $F\,\Delta$ |
| (Constant) | 1.620 | .296 | | 5.475*** | | | |
| Block 1: Demographics | | | | | .001 | .001 | 0.26 |
| Gender | -.083 | .055 | -.036 | -1.497 | | | |
| First generation | -.033 | .049 | -.016 | -0.677 | | | |
| International student | -.030 | .076 | -.010 | -0.392 | | | |
| Childhood language | -.149 | .057 | -.068 | -2.601** | | | |
| Block 2: Universities | | | | | .006 | .005 | 2.43 |
| Waterloo | -.042 | .085 | -.013 | -0.494 | | | |
| Western | -.183 | .096 | -.049 | -1.907 | | | |
| Toronto | -.007 | .058 | -.003 | -0.127 | | | |
| | Estimates | | | | Blocks | | |

| Model 5 | *b* | *Std Error* | *Beta* | *t* | $R^2$ | $\Delta R^2$ | $F \Delta$ |
|---|---|---|---|---|---|---|---|
| Block 3: Year level | | | | | .015 | .009 | 3.39** |
| Year 2 | .215 | .066 | .089 | 3.229** | | | |
| Year 3 | .291 | .071 | .113 | 4.124*** | | | |
| Year 4 | .294 | .080 | .098 | 3.663*** | | | |
| Year 4+ | .445 | .084 | .141 | 5.304*** | | | |
| Block 4: Grades | | | | | .117 | .102 | 84.83*** |
| High school | .087 | .022 | .100 | 3.880*** | | | |
| University | -.167 | .020 | -.224 | -8.332*** | | | |
| Block 5: Academic capital | | | | | .195 | .078 | 70.66*** |
| Abstract deficiency | .209 | .054 | .132 | 3.847*** | | | |
| Concrete deficiency | .308 | .052 | .204 | 5.903*** | | | |

*Note*. Listwise deletion, $N = 1467$; * $p < .05$; ** $p < .01$. *** $p < .001$.

As with thoughts of dropping out, each year level represented slightly increased dissatisfaction over first year, but most of the variance in the final model was explained by grades and the measures of academic capital deficiencies. Grades explained the most variance ($R^2$ change = .102), followed by the two academic capital deficiency scales ($R^2$ = .078). Looking at the estimates for individual variables, it can be seen that university grades had the largest beta (-.224), followed by concrete (.204) and then abstract (.132) academic capital deficiency; however, the combined betas for capital (.336) exceed that of university grades (-.224). In addition, as with thoughts of dropping out, high school grades were unexpectedly positively associated with dissatisfaction (i.e., the higher their grades coming out of high school, the more dissatisfied students were of their university experience, again pointing to the possible negative effects of high school grade inflation).

As we did with the analysis of thoughts of dropping out, when the skill-proficiency clusters were substituted for the two academic capital scales in a regression analysis with the same independent variables entered in the same order, roughly the same amount of variance was explained by the overall model ($R^2$ = .161), and the clusters uniquely explained roughly the same amount of variance ($R^2$ = .054) as the two scales. The beta for the at-risk group was .160, and for the dysfunctional group, it was .249. Again, either estimate of academic capital deficiency can be used in future studies, depending on the purpose of the investigation.

## Discussion

This study found that seven factors of academic capital could be reliably measured, six of which created a two-factor solution measuring abstract and concrete deficiencies. In addition to this evidence of factorial validity, evidence for predictive criterion validity (Anastasi, 1988) of these measures was found in terms of hypothesized associations with student grades, satisfaction, and thoughts of dropping out. In addition to the measurement scales, three clusters of skill-proficiency were identified. A majority of the university students in the sample were deemed at-risk of sub-optimal academic outcomes, including a sizeable minority (15%) who scored as dysfunctional in the university context.

When applied across four of Ontario's largest universities, it was found that this level of risk and dysfunction does not differ substantially among those universities, in spite of their significantly different entrance standards in terms of incoming high school grades. Moreover, the proportions of at-risk and dysfunctional students did not meaningfully decrease with year level. Controlling for other variables, demographic variables showed only minor associations with skill deficits, and first-generation student status was unrelated to these deficiencies.

The overall pattern of results points to significant grade inflation at both the high school and university levels, as well as limited skills development while students are attending university. Interestingly,

high school grades explain very little in the bivariate analyses and are uncorrelated with dissatisfaction with the university experience or increased thoughts of dropping out. Yet, in the regressions, both of these negative experiences are positively associated with high-school grades—controlling for the other variables, including academic capital deficiencies, the analysis reveals a suppressed relationship such that students sent to university with high grades, but low skills, are more likely to encounter these negative experiences.

The results also suggest that high school grade inflation may be a more pressing concern at universities that rely on very high incoming grades for admissions, such as Western. Western has the highest incoming grades of the four universities but also the largest grade drop. In contrast, York has the most functional students even though it has the lowest level of incoming high school grades (B+) and the lowest grade drop. Given that the highest graded incoming students may not have the commensurately highest skill levels, the use of high school grades in the A/A+ range as a selection method may need to be reviewed. Indeed, the implications of these findings are serious from a policy perspective, given that Western's official policy has been to "attract the brightest students as demonstrated through the highest entering grade average" (http://www.ipb.uwo.ca/pi/secondary_school_grades_of_incoming_students.php). Yet, our results (Table 3) suggest that the academic skill levels of Western students are indistinguishable from the other three universities, including York students who have substantially lower incoming grades.

## Policy Implications

The spectre of grade inflation is seen in the finding that dysfunctional students receive relatively good grades in high school (B+) and "satisfactory" grades in university (C+). As a result, the question to be answered at the policy level is: "how are these grades possible among students who lack basic skills?"

The possibility that students are receiving inflated grades at both levels to keep them in school needs to become a focus of future policy research, as should the association of deficiencies in academic capital levels with two negative student experiences—dissatisfaction with the university experience and thoughts of dropping out. Both of these are important indicators that educators, administrators, and policymakers need to take seriously in terms of both the human (student stress) and financial (lost tuition fees, misdirected teaching and counselling resources, costs to the economy, etc.) costs. Our research suggests that improving students' academic skill levels can reduce these negative experiences.

As noted in the introduction, one of the objectives of Ontario's Ministry of Education (2010) is to facilitate the development of various skills and work habits to prepare students for postsecondary education and the workforce. Despite this objective, many employers, academics, and research agencies believe that Ontario's high schools are not teaching skills consistent with Ministry objectives. Indeed, as suggested by our research and that of HEQCO, not only are undergraduates' skill levels relatively low as a group, but also they enter and leave university with little or no improvement.

While HEQCO has focused on important abstract skills, they have not considered the day-to-day abilities that would assist undergraduates in their encounters with the demands of academic life, nor have they shown an association between the skills they measure and academic success. For this reason, in this article, we focused on these lacunae. Our approach was distinct from that of HEQCO in that we utilized the concept of academic capital while also capturing the experiences of students via their own assessments. At the same time, consistent with HEQCO and Statistics Canada's research, our endeavour suggested that large numbers of students leave Ontario's high schools lacking the academic capital associated with university and employment success. Moreover, students acquire no appreciable increases in this important asset over the course of their university careers.

One aspect of this problem can be clearly understood in terms of the current grading practices that appear to obviate the relationship between the Ministry of Education's laudable objectives and the process whereby it measures the attainment of those goals. According to the Ministry, "evaluation accurately summarizes and communicates to parents, other teachers, employers, institutions of further education, and students themselves what students know and can do with respect to the overall curriculum expectations" (Ministry of Education, 2010, p. 38). Such evaluations are ostensibly based on measures of students' "achievement of curriculum expectations, and … *the development of learning skills and work habits*" (our emphasis) (Ministry of Education, 2010, p. 44). Evaluations, themselves, are based on the

results of tests and assignments, performance tasks, demonstrations, projects, essays, and group projects (Ministry of Education, 2010, p. 39).

Whatever the measure employed in student evaluation, individual teachers known to students, parents, and principals are currently responsible for assigning grades. The Ministry expects that in their evaluations of tasks, teachers "will benefit from leadership by the principal to ensure that there is a common understanding among all staff about the process for determining the final grade" (Ministry of Education, 2010, p. 39). This said, in its discussion of assessment, the Ministry gives a candid recognition of the limitations of the process:

> It is worth noting, right from the start, that assessment is a human process, conducted by and with human beings, and subject inevitably to the frailties of human judgement. However crisp and objective we might try to make it, and however neatly quantifiable may be our 'results,' assessment is closer to an art than a science. It is, after all, an exercise in human communication (Ministry of Education, 2010, p. 29).

Consistent with the above recognition, it is clear that even with the best of intentions, assessments would vary from one teacher to the next. This human limitation can be accepted. Less acceptable is any mistakable belief that in their determinations teachers are immune from the criticisms of students, parents, others employed in the educational system, and school trustees. In this type of environment, teachers, rather than students themselves, may be held accountable for low grades (e.g., they face claims that they don't know how to teach). When confronted by possibilities such as these, it is understandable that certain teachers would award undeserved or inflated grades. However, we wish to make it clear that we believe the blame for the skill problems we have identified should not be laid at the feet of secondary school teachers or university professors. These problems are part of a systemic product of a secondary school system with praiseworthy goals that lack the means to assess the achievement of these goals, and a post-secondary system in which penalties may be incurred by those who hold students accountable for important academic skills.

In Alberta, it is relatively easy to determine the degree of grade inflation likely resulting from pressures such as the foregoing. In that province, home teachers assign 70% of students' final high school grades. The other 30% is measured through province-wide tests, the marking of which involves no home teachers. Overall, in 2018, home teacher marked tests were 8% higher than those graded at the provincial level (Alberta Government, 2019). At the university level, this difference represents, for example, the difference between a B and a C. It is important to note that tests conducted at the provincial level can contain both multiple-choice and essay-type questions. We mention this lest readers assume that differences in grading are simply a function of the ways in which questions are asked.

Unfortunately, information similar to that collected in Alberta is unavailable for Ontario. However, until 1968, entrance to university in Ontario was based exclusively on the final examination results of students' last year of high school (grade 13) (Allen et al., 1983, p. 38). These were not standardized multiple-choice tests. Instead, particularly in the disciplines now considered part of the social sciences and humanities, students responded to essay-type questions in three hour examinations. These were primarily of the essay or short-answer type and graded in a central location in Toronto. In this situation, any high school politics were minimized, a change that over the past few decades may have led to grade inflation and inconsistencies in standards from one high school to the next (Côté & Allahar, 2007, 2011).

After the abolition of centrally marked examinations, the situation changed immediately and drastically. As shown by Allen et al. (1983, p. 49), in 1965-66, 6.6% of Ontario's grade 13 graduates earned first class honours. By 1979-80 the figure had risen to 25.3%. In other words, the transition to a system in which assessments were conducted in students' home schools was consistent with grade inflation. Indeed, as early as 1970, university administrators, because of grade inflation, were lamenting the passing of the old grade 13 system (Saywell, 2008). Grade inflation has accelerated in the intervening years, with over half of university applicants presenting high school grades of A or A+ (e.g., Côté & Allahar, 2007).

The same trend was evident within universities. An examination of grades conducted at seven Ontario universities showed that between the mid-seventies and mid-nineties, grades inflated by 5% (Anglin & Meng, 2000). Similarly, grades inflated at Western in the mid-1990s in conjunction with the introduction of "Enrollment Contingent Funding" (known as "bums on seats" funding), with As and Bs increasing by 10-20% in the Faculties at Social Science, Arts & Humanities, and Natural Science (Côté

& Allahar, 2010). Elsewhere, a far-reaching analysis found significant grade inflation in all classifications of American universities (Rojstaczer, 2016), and it is increasingly being acknowledged as a world-wide phenomenon (e.g., Baker, 2018).

In view of these considerations, the first step toward ensuring that students leaving high school in Ontario have the forms of academic capital embodied in the Ministry's objectives would be the introduction of province-wide tests and arms-length evaluation. The tests should not be standardized multiple-choice examinations. Instead, where possible, they should be constructed in a way such that carefully trained graders would be able to identify students' thinking patterns as revealed in responses to essay-type or short-answer questions. In other words, students would not simply be right or wrong. Practices of this nature would not penalize, for example, culturally disadvantaged students to the same extent as standardized multiple-choice tests apparently do.

While a system of this nature would be expensive, future savings resulting from better assessments of students' academic attainment would outweigh the costs: there would be less of a gap between high school and university grades; universities would spend less on providing remedial help; students would be more likely than now to complete their degrees; employers would spend less on taking steps to bring new employees up to acceptable standards. As an interim measure, universities should introduce courses focusing on the development of forms of academic capital consistent with academic success. Details of one such initiative can be found at https://skillsforuniversitysuccess.info.yorku.ca.

## Limitations and Future Research

The present study was unfunded and therefore limited in a number of ways in terms of data collection. It was cross-sectional when the better methodology is longitudinal. Its sampling was based on email lists provided by university registrars, so students were approached without any lead publicity or incentives for participation. Clearly, larger random samples are preferred in which students are followed longitudinally, with data culled from multiple sources: institutional data, such as grades; ratings of student skills by independent observers; and self-reported data as appropriate for more subjective issues like satisfaction.

Given that the costs of such research run into the millions of dollars and data collection requires close cooperation by universities, we call upon universities themselves to take up the types of measures we have developed for use as diagnostic "screens." They could be included with other sources of data collection. As agencies like HEQCO and Statistics Canada have suitable resources, they could take the next step and find more definitive answers to the questions we have raised.

## References

Alberta Government. (2019). *Diploma Examination Multiyear Reports.* https://education.alberta.ca/media/3772499/diploma-multiyear-province-report-table.pdf

Allen, L. G., Alexander, L. D., Hughes, R. C., & Rosenfeld, J. M. (1983). An examination of performance of first year students at an Onatrio university: An admission perspective. *The Canadian Journal of Higher Education, 13*(3), 37–54.

Anastasi, A. (1988). *Psychological testing* (6th ed.). MacMillan.

Anglin, P. M., & Meng, R. (2000). Evidence on grades and grade inflation at Ontario's universities. *Canadian Public Policy, 26*, 361–368.

Arum, R., & Roska, J. (2011). *Academically adrift: Limited learning on college campuses.* University of Chicago Press.

Baker, S. (2018, June 28). Is grade inflation a world-wide trend? *Times Higher Education.* https://www.timeshighereducation.com/features/grade-inflation-worldwide-trend

Bradshaw, J. (2011, May 9). When a university degree just isn't enough. *Globe and Mail.* https://www.theglobeandmail.com/news/national/time-to-lead/when-a-university-degree-just-isnt-enough/article579230/

Brown, L. (2016, February 22). Young grads need to brush up on 3 Rs, employers say. *The Star.*

Cain, P. (2018, September 13). *One university's secret list to judge applicants by their high schools - not just their marks.* https://globalnews.ca/news/4405495/waterloo-engineering-grade-inflation-list/:

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Côté, J. E., & Allahar, A. L. (2007). *Ivory tower blues: A university system in crisis.* University of Toronto Press.

Côté, J. E., & Allahar, A. L. (2010, March 25). Grade inflation: A tale of three Western faculites. *Western News,* 5.

Côté, J. E., & Allahar, A. L. (2011). *Lowering higher education: The rise of corporate universities and the fall of liberal education.* University of Toronto Press.

Dion, N., & Maldonado, V. (2013, December 4). We need to assess student literacy skills. *University Affairs.*

Dwyer, M. (2017, November 13). What grades do I need to get into Canadian univesities? *MacLean's.* https://www.macleans.ca/education/what-grades-do-i-need-to-get-into-canadian-universities/

HEQCO. (2017). *HEQCO Essential Adult Skills Initiative 2016/2017: Status Report.* http://www.heqco.ca/SiteCollectionDocuments/HEQCO%20EASI_Status%20Report%20on%20College%20Pilot_March2017_English.pdf:

Hutchins, A. (2017, April 17). Which universities prepare students for employment? *Macleans.*

Kuncel, N. R., Credé, M., & Thomas, L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75*, 63–82.

Laurie, R. (2007). *Setting them up to fail? Excellent school marks don't necessarily lead to excellent exam marks.* Atlantic Institute of Market Studies. http://www.aims.ca/site/media/aims/Gradeinflationvfinalcc.pdf.

Levine, A. & Cureton, J. S. (1998). *When hope and fear collide: A portrait of today's college students.* Jossey Bass.

Liu, O. L., Mao, L., Frankel, L., & Xu, J. (2016). Assessing critical thinking in higher education: The HEIghten approach and preliminary validity evidence. *Assessment and Evaluation in Higher Education, 41*(5), 677-694.

Mahboui, P. (2017). *Talkin' Bout my generation: More education less skilled Canadians.* https://www.cdhowe.org/public-policy-research/talkin%E2%80%99-%E2%80%98bout-my-generation-more-educated-less-skilled-canadians.

Mancuso, M. (2014, January 15). Which skills do our graduates lack? *University Affairs.*

Ministry of Education. (2010). *Growing success: Assessment, evaluation, and reporting in Ontario schools.* http://www.edu.gov.on.ca/eng/policyfunding/growSuccess.pdf

Ministry of Education. (2018). *The Ontario curriculum grades 11 and 12: Cooperative education.* https://www.ontario.ca/page/ministry-education

Munro, D. (2014). *Skills and higher education in Canada.* Statistics Canada.

Rojstaczer, S. (2016). *Grade inflation at American colleges and universities.* http://www.gradeinflation.com

Saywell, J. T. (2008). *Someone to teach them: York and the great university explosion, 1960-1973.* University of Toronto Press.

Statistics Canada. (2016, September 14). Overqualification, skills and job satisfaction. *The Daily.*

Tsui, L. (2002). Fostering critical thinking through effective pedagogy: Evidence from four institutional case studies. *The Journal of Higher Education, 73*(6), 740-763.

Weingarten, H. P. (2014, September 12). We have to measure literacy and numeracy among university graduates. *Globe and Mail.*

Weingarten, H. P., & Hicks, M. (2018). *On test: Skills, summary of findings from HECQO's skills assessment pilot studies.* http://www.heqco.ca/en-ca/Research/ResPub/Pages/On-Test-Skills-Summary-of-Findings-from-HEQCO's-Skills-Assessment-Pilot-Studies.aspx

Williams, R. L., & Stockade, S. L. (2003). High-performing students with low critical thinking skills. *The Journal of General Education, 52*(3), 200-226.

## Appendix A

Students' Skill Deficiencies on each Item for the Seven Subscales

| % Difficult/Very Difficult or Unconfident/Very Unconfident | Toronto | Waterloo | Western | York | Overall |
|---|---|---|---|---|---|
| ***Writing*** | | | | | |
| Difficulty with one-page summary of article | 35% | 26% | 28% | 28% | 30% |
| Difficulty with formulating theses for essays | 26% | 14% | 29% | 28% | 26% |
| Lack of confidence in ability to apply APA/Chicago style in written work | 23% | 32% | 27% | 25% | 25% |
| Lack of confidence in knowing sequencing in APA/Chicago | 43% | 43% | 42% | 38% | 41% |
| Difficulty with evaluating appropriateness of referencing in essays | 23% | 26% | 25% | 27% | 25% |
| Difficulty with evaluating appropriateness of footnote usage in essays | 38% | 49% | 39% | 38% | 39% |
| Lack of confidence in own use of English grammar | 15% | 11% | 18% | 19% | 17% |
| Difficulty with identifying grammatical mistakes | 26% | 21% | 22% | 29% | 26% |
| Difficulty with identifying plagiarism in essays | 21% | 21% | 25% | 21% | 22% |
| Difficulty with summary of thesis and strengths of article | 23% | 18% | 24% | 21% | 21% |
| Lack of confidence in knowledge of writing critical book review | 49% | 54% | 53% | 48% | 49% |
| Lack of confidence in writing effective papers | 15% | 11% | 23% | 20% | 18% |
| Difficulty with taking effective lecture notes | 25% | 32% | 40% | 30% | 29% |
| *Mean* | 28% | 28% | 30% | 29% | 28% |
| ***Test taking*** | | | | | |
| Difficulty with studying for tests | 36% | 32% | 39% | 32% | 34% |
| Lack of confidence in approach to short test questions | 24% | 22% | 27% | 26% | 25% |

| % Difficult/Very Difficult or Unconfident/Very Unconfident | Toronto | Waterloo | Western | York | Overall |
|---|---|---|---|---|---|
| Lack of confidence in approach to long test questions | 32% | 28% | 34% | 34% | 33% |
| Lack of confidence in approach to multiple-choice test questions | 20% | 15% | 23% | 17% | 18% |
| *Mean* | 28% | 24% | 31% | 27% | 27% |
| ***Analysis*** | | | | | |
| Difficulty with determining main point in articles | 30% | 28% | 24% | 27% | 27% |
| Difficulty with determining perspective of articles | 26% | 17% | 26% | 23% | 23% |
| Difficulty with synthesizing views of articles | 24% | 19% | 25% | 25% | 24% |
| Difficulty with identifying types of argument in articles | 57% | 60% | 57% | 56% | 57% |
| Difficulty with determining bias and its effect on argument in articles | 26% | 17% | 26% | 25% | 24% |
| Difficulty with interpreting numerical tables | 20% | 15% | 18% | 13% | 16% |
| Difficulty with integrating new knowledge into moral framework | 25% | 22% | 21% | 22% | 22% |
| *Mean* | 29% | 25% | 28% | 27% | 28% |
| ***Time and Group Management*** | | | | | |
| Lack of confidence in time management skills | 32% | 29% | 35% | 28% | 30% |
| Lack of confidence in study-group management skills | 20% | 17% | 21% | 20% | 20% |
| Lack of confidence can foster group productivity | 27% | 31% | 26% | 23% | 25% |
| Lack of confidence in handling problem group members | 25% | 26% | 29% | 20% | 23% |
| *Mean* | 26% | 26% | 28% | 23% | 24% |
| ***Research*** | | | | | |
| Difficulty with identifying good evidence | 15% | 8% | 12% | 15% | 14% |

| % Difficult/Very Difficult or Unconfident/Very Unconfident | Toronto | Waterloo | Western | York | Overall |
|---|---|---|---|---|---|
| Difficulty with identifying good sources for essays | 35% | 31% | 27% | 27% | 29% |
| Lack of confidence in online search abilities for essays | 15% | 8% | 12% | 15% | 14% |
| Lack of confidence in other than online search abilities for essays | 35% | 31% | 27% | 27% | 29% |
| Lack of confidence in identifying scholarly sources | 15% | 13% | 17% | 21% | 18% |
| *Mean* | 22% | 18% | 19% | 21% | 21% |
| ***Presentations*** | | | | | |
| Difficulty with verbally summarizing articles' strengths and weaknesses | 25% | 16% | 22% | 22% | 22% |
| Difficulty with preparing for presentations | 25% | 16% | 22% | 22% | 22% |
| Difficulty with formulating argument for presentations | 19% | 12% | 16% | 20% | 19% |
| Difficulty with giving presentations | 48% | 32% | 34% | 36% | 37% |
| Difficulty with verbally handling criticism in presentations | 35% | 31% | 38% | 31% | 32% |
| Difficulty with using technological enhancements for presentations | 7% | 3% | 6% | 9% | 7% |
| Difficulty with speaking clearly when giving presentations | 27% | 15% | 27% | 20% | 22% |
| Difficulty with identifying controversial issues in presentations of others | 15% | 7% | 14% | 14% | 13% |
| Difficulty with identifying flaws in presentations of others | 26% | 23% | 32% | 22% | 24% |
| *Mean* | 24% | 17% | 23% | 22% | 22% |
| ***Numeracy*** | | | | | |
| Difficulty with elementary subtraction | 8% | 8% | 8% | 7% | 8% |
| Difficulty with elementary percentage calculation | 33% | 42% | 39% | 26% | 31% |
| Difficulty with elementary multiplication | 26% | 37% | 32% | 23% | 26% |

| % Difficult/Very Difficult or Unconfident/Very Unconfident | Toronto | Waterloo | Western | York | Overall |
|---|---|---|---|---|---|
| Difficulty with elementary addition | 8% | 7% | 7% | 9% | 8% |
| Difficulty with elementary algebra | 18% | 28% | 21% | 26% | 24% |
| Difficulty with explanation of square root | 19% | 26% | 19% | 27% | 24% |
| *Mean* | 18% | 25% | 21% | 20% | 20% |