WHY THE RANKINGS ARE WRONG: LESSONS FROM DATA UNDERLYING U.K. UNIVERSITY LEAGUE TABLES

J. PAUL GRAYSON
YORK UNIVERSITY

KYLE GRAYSON
NEWCASTLE UNIVERSITY

Abstract

This article uses high-quality data from the United Kingdom to critically examine the assumptions underpinning university league tables in Canada and elsewhere. Particular attention is given to research quality and value added in learning (VA). Unlike many global ranking systems, U.K. rankings incorporate measures that allow for robust assessment. Based on these data, we show that league table positions in the United Kingdom are weakly associated with research quality and largely unrelated to VA. Furthermore, research funding bears little consistent relationship to research quality. At both institutional and departmental levels, notably in Politics and International Studies, highly ranked universities do not necessarily deliver superior learning outcomes or research quality. These findings challenge the validity of using inputs such as research grants or student awards as proxies for institutional quality. Although U.K. data cannot be generalized automatically, they provide a valuable reference point for other countries. As a result of these findings, we argue that Canadian rankings, and others like them, must be treated with considerable skepticism.
Keywords: university league tables, Canada, United Kingdom, validity university rankings

Résumé

Le présent article utilise des données britanniques de grande qualité pour examiner de manière critique les présupposés qui sous-tendent les classements universitaires au Canada et ailleurs dans le monde. Une attention particulière est accordée à la qualité de la recherche et à la valeur ajoutée dans l'apprentissage. Contrairement à de nombreux systèmes de classement mondiaux, les classements britanniques intègrent des indicateurs qui permettent une évaluation robuste. Sur la base de ces données, nous montrons que les positions dans les classements au Royaume-Uni sont faiblement associées à la qualité de la recherche et n'ont pratiquement aucun lien avec la valeur ajoutée. De plus, le rapport entre le financement de la recherche et la qualité de la recherche est peu constant. Aux niveaux institutionnel et départemental, notamment en sciences politiques et études internationales, les universités les mieux classées ne sont pas nécessairement supérieures en matière de résultats d'apprentissage ou de qualité de la recherche. Ces résultats remettent en question la validité de l'utilisation de données telles que les subventions de recherche ou les bourses d'études comme indicateurs de la qualité institutionnelle. Bien que les données

du Royaume-Uni ne puissent pas être généralisées automatiquement, elles constituent un point de repère précieux pour les autres pays. En conséquence, nous soutenons que les classements canadiens, et d'autres classements semblables, doivent être considérés avec beaucoup de scepticisme.
Mots clés : enseignement supérieur canadien, comportement de recherche d'aide, étudiants internationaux, soutien social

INTRODUCTION

Every year, *Maclean's* magazine collects information on Canadian universities. These data are then used to give each university a score. Based on these scores, universities are arranged hierarchically in what are known in the United Kingdom as "league tables." Such rankings can be used to lure students, impress donors, and woo governments.

For its league tables, *Maclean's* obtains and utilizes information in a number of ways. For example, in its calculations, it includes national academic awards won by students and faculty, the value of faculty research grants, the size of a university's operating budget, spending on student services, and university reputation as determined by professors, administrators, and businesspeople (*Maclean's*, 2023). In each case, more is better.

Unfortunately, *Maclean's* provides little information on the concepts these measures presumably operationalize and their possible connections with one another. Are we to assume that awards won by students reflect the university's impact on learning? What about faculty awards? Do they indicate research quality? Are research grants also a reflection of research quality? Is there a connection between research quality and how much students learn? From the way *Maclean's* summarizes its findings, we do not know the answers to the above. All we do know is that values on each of several variables find expression in league tables.

In contrast to other nations, the United Kingdom collects more precise information on a number of variables, including value added in learning and the quality of faculty research. In this article we utilize these data to examine some unanswered questions deriving from *Maclean's* magazine's use of certain variables in their rankings. In other words, our concern is methodological, not theoretical or normative: rankings are not going away, how might their underlying variables be selected to better reflect what they claim to be measuring? Thus, in this article, we are engaged in an immanent critique that evaluates the practices of *Maclean's* based on their own self-understanding of what their ranking practices measure.

More specifically, at the institutional level, what are the relationships among research quality, university ranking, and value added in learning? For universities and Departments of Political and International Studies in the United Kingdom, we will show that there are small, or no, relationships among these variables. While the U.K. university system differs from those elsewhere, similar links can be hypothesized for other nations, including Canada. Overall, we will show that many measures used by *Maclean's* to operationalize the constructs implicit in their ranking of universities are questionable.

THE CONTEXT

Rankings like *Maclean's* are common. As shown elsewhere (Grayson & Grayson, 2023), at both the national and international levels, organizations make attempts to assess the quality of universities (Moed, 2017) in similar ways. Various national and international agencies, in a mainly atheoretical fashion, collect readily available information (Brankovic, 2021; Harvey, 2008; Lynch, 2015). In the absence of a clear definition of quality, and with little apparent consideration

for validity (Dill & Beerkens, 2005), this information may include a university's average class size, money spent on extracurricular activities, availability of on-campus accommodation, the ratio of international to domestic students, and the magnitude of alumni donations. Algorithms are then applied to the data resulting in an overall institutional score. These scores are then expressed in league tables ranging from high to low. The higher the score, the greater the presumed quality of a university. Under these circumstances, "university rankings have become a proxy for quality and excellence" (Easley et al., 2021, p. 306).

Independent of provenance, all rankings share some common limitations. It is not the intent to adumbrate each. Suffice it to say that—among other issues—many concepts are poorly explained; insufficient information is provided on methodology; the goals of individual universities are ignored; operationalizations of many variables lack validity; some data are of poor quality; little consideration is given to the constraints faced by universities in the Global South; key variables often reflect a "neo-liberal" understanding of the purposes of higher education; rankings ignore that higher education providers occupy niches in [inter]national ecosystems; and there is an assumption of a qualitative hierarchy of institutions (Amsler & Bolsmann, 2012; Bowden, 2000; Calderon, 2021; David, 2016; Foley & Goldstein, 2012; Gruber, 2014; Jarocka, 2012; Jung et al., 2011; Ordorika & Lloyd, 2014).

Despite their limitations, rankings result in national and international status systems in which universities compete for position (Marginson, 2006). Rankings are thus closely associated with market-based approaches to higher education (Amsler & Bolsmann, 2012; Ordorika & Lloyd, 2014). Within this framework, education is a "product." Students are "consumers" (McManus et al., 2017).

At the international level, universities are monitored by organizations such as Quacquarelli Symonds (QS), The Times Educational World University Rankings (Schleicher, 2015), and the Academic Ranking of World Universities (ARWU). Every year they compare post-secondary institutions in the United Kingdom, the United States, China, France, and elsewhere. Others, like *Maclean's*, conduct within-country comparisons. The Center for Higher Education (CHE) University Ranking and La Republica are included in this category. The former focuses on Germany and Austria; the latter on Italy.

Despite their widespread use, information needed to answer questions regarding the validity of measures employed in university rankings is hard to find. For example, to determine the amount of learning that takes place on a campus, we must have a measure of students' abilities at two points in time: upon enrolment and graduation (Mayhew et al., 2016; Pascarella & Terenzini, 1991, 2005). The difference between these two can be defined as "value added in learning" (or simply value added; VA). In the absence of this information, we will not know if achievement of graduates, as manifested in the winning of external scholarships, etc., is a product of university experiences or the quality of recruited students. Such information is unavailable for Canada and most other countries.

Defining research quality is difficult. Quite simply, the characteristics of excellent research vary from one discipline to the next. As a result, it is best to let those in any given field determine quality (Smith, 1993).

Peer review of research products is one way of reaching this goal. Ideally, in conducting their assessments, reviewers employ common standards. In the quantitative social sciences, of which this article is an example, such standards may include: the nature of the research design, the methodology, data quality and analysis, and the extent to which results are valid, reliable, ethical, and useful. For other disciplines, and for other methodological approaches, criteria could be different. Quality may also be determined by, for example, the number of times a work in a given field is cited. There are strengths and weaknesses associated with these and other approaches (this will be explored further later in the article).

Consistent with the foregoing, it may be misleading to assume that the receipt of research grants is a measure of research quality. A research grant is an input, not an output.

More important is the association between research grants and the quality of research resulting therefrom. Do those who receive grants conduct better research? Again, access to Canadian and other information that is needed to provide a clear answer is not readily available.

In view of the foregoing nuances, it is not surprising that many organizations producing league tables manifest shortcomings like those of *Maclean's* (David, 2016; Dolton & Makepeace, 1982; Gruber, 2014; Harvey, 2008; Jung et al., 2011; Marginson, 2006; Ordorika & Lloyd, 2014). This is certainly the case in the United Kingdom, where league tables are produced by the *Complete University Guide* (CUG), *The Guardian*, and *The Times* (Amsler & Bolsmann, 2012; Boliver, 2015; Broecke, 2015).

Where these differ from *Maclean's* is in access to, and use of, some high-quality data. These data include national information that contributes to calculations of value added in learning and research quality. Each is part of a university's *raison d'être*. Indeed, our position is that in combination these constitute key objectives of post-secondary practice.

Although, like *Maclean's*, the *CUG* and *The Guardian* fall short in data utilization, analyses of the high-quality information at their disposal is of potential relevance to the interpretation of Canadian and other league tables. More specifically, it is possible to use U.K. data to provide answers to some of the questions raised above. These include unarticulated yet implicit assumptions regarding student and university characteristics conducive to learning, and the relationship between research funding and research quality.

When examining these issues, it is essential to acknowledge that the U.K. university system is unique. Nonetheless, *faute de mieux*, an examination of data from the United Kingdom results in several, possibly universally relevant, insights. Unfortunately, while we benefit from the insights, the probability that in Canada, and likely elsewhere, data will ever be available to test hypotheses deriving therefrom is low. As currently constituted, the Canadian and other educational systems are unable to provide the required information.

Why does any of this matter? As shown elsewhere (Grayson & Grayson, 2023), among other reasons, despite their glaring limitations, some students take league tables seriously. Research conducted in the United States (Zilvinskis & Rocconi, 2018, p. 257), Canada (Drewes & Michael, 2006, p. 799), and the United Kingdom (Broecke, 2015; Chevalier & Jia, 2015; Gibbons et al., 2015) shows that a university's standing may have an influence, albeit small, on numbers of university applications. More importantly, university rankings constitute a status system in which all universities feel the need to compete. In so doing they may exaggerate positive—and ignore negative—characteristics. These practices are to the detriment of potential students, who would benefit from a better understanding of what an institution can offer them.

THE CONSTRUCTION OF U.K. LEAGUE TABLES

In the United Kingdom, as shown elsewhere (Grayson & Grayson, 2023), there are prominent media organizations that, on an annual basis, construct league tables. In conducting their activities, the rankers utilize information collected by the National Student Survey (NSS; 2021), the Higher Education Statistics Agency (HESA), and the Research Excellence Framework (REF; 2021). *The Times* supplements this material with assessments of the reputations of various universities.

The NSS and HESA

The National Student Survey is made available to all students in their final year of undergraduate study. In 2019, the last pre-pandemic year, the response rate was an impressive 72% of those in their last year of university. This figure represents 339,000 students (NSS, 2020, p. 1).

Perhaps most importantly, the survey asks students to comment on various aspects of their university lives. Unfortunately, not all institutions participate in this endeavour and some fall below response rate thresholds. As a result, for example, satisfaction levels for students at Oxford

and Cambridge have often been unavailable. While the *CUG* accepts the absence of such information, *The Guardian* imputes values for missing data (Hiely-Rayner, 2016). Despite the strengths of the National Student Survey, it cannot measure the satisfaction (or dissatisfaction) of students leaving prior to degree completion, and the potential role of satisfaction in their exit.

Primarily from administrative records, the Higher Education Statistics Agency collects information on a wide range of university characteristics such as entry standards and graduation rates.

While information collected by the National Student Survey and the Higher Education Statistics Agency is robust, reputational data collected by *The Times* are different. For some scholars, rankings should be based in reality, not perceptions of reality as embodied in reputational surveys (Marginson, 2014). By definition, the validity and reliability of reputational measures cannot be determined (Soysal et al., 2020). In view of the foregoing, the utilization of reputational data in university ranking systems, as in other areas of social science, is problematic. For these reasons, we focus on the rankings provided by the *CUG* and *The Guardian*.

Recall that *Maclean's* also includes reputational surveys in its ranking of Canadian universities. The problems with their inclusion are identical to those identified in connection with the *The Times*. This is yet another reason to view *Maclean's* rankings with caution.

Value Added in Learning

Although the concept is not new (Chapman, 1996; Schleicher, 2015), *The Guardian* has developed a very helpful measure of value added (Hiely-Rayner, 2016). Via this calculation it aims to show how much knowledge and skills students have acquired over the course of their studies.

The Guardian operationalizes its concept of value added in the following manner. Based on entering grades, it predicts the probability of a student achieving a "good degree" (a first- or upper second-class in the U.K. system) upon graduation. In most cases, entry grades are the results of national A-level examinations marked by evaluators external to students' schools. The value added is the difference between the predicted and the actual number earning a good degree several years later. As scored by *The Guardian*, value added can range from 1 to 10. The higher the number, the greater the value added.

Quality of Research: The Research Excellence Framework

Approximately every seven years, under the auspices of the Research Excellence Framework (REF, 2014, 2021), upon request, exhaustive reviews are conducted of the research of U.K. university subject groups. These groups may reflect department-specific disciplines or areas of inquiry that span departments. The fact that units applying receive government research funding based on their evaluation is an incentive to participate.

The most recent seven-year REF was conducted in 2021 (REF, 2021). We used this information sparingly in the current analysis for two reasons. First, most of the data for this and an associated project (Grayson & Grayson, 2023) were collected prior to the 2022 release of REF 2021. Second, in the associated project REF scores were related to student outcomes potentially affected by the pandemic. Accordingly, in much of the current analysis we felt it best to focus mainly on information in the pre-pandemic league tables of 2019. The REF utilized in that year was conducted in 2014. This said, for the analysis of a particular discipline, Politics and International Studies, we utilized the REF 2021.

Research Excellence Framework 2014 (and 2021) involved an external assessment of research submitted by subject areas across the United Kingdom by independent committees of experts. In 2014, 65% of a total score was based on research outputs. A further 15% derived from the nature of the departmental research environment, and 20% from the societal impact of research. Slight changes were made for 2021 to decrease the weighting assigned to outputs (60%) and increase the weighting of societal impact (25%).

All outputs considered in the REF, from individual research papers to impact case studies,

were assessed on a scale from 0–4*. Zero represented "quality that falls below the standard of nationally recognised work…or work which does not meet the published definition of research for the purposes of this assessment" and 4* signified "quality that is world-leading in terms of originality, significance and rigour" (REF, 2014, 2021). All distributions of scores for subjects (called units of assessment) submitted by a university were aggregated to provide tallies at the institutional level. In any institution it was possible that not all departments engaged in the REF process.

In 2018–19, the *CUG* presented the results of the REF 2014 along two dimensions: research quality and research intensity. To make use of this information we gave equal weight to each: (((Research Quality/4) + (Research Intensity))/2)*100). Using this formula, the maximum score obtained by a university was 89%. This figure does not preclude the possibility that a subject area may have received a perfect score of 100%.

It is important to emphasize that the REF focuses exclusively on the quality of research. It is uninterested in the status of journals in which research appears. It has no interest in the number of times a work is referenced (that may simply reflect numbers of researchers in a field of inquiry) or how much it received in support via research grants and contracts. This is important. Each may sometimes, and unwarrantedly, be used as a proxy for quality itself. When this occurs, superior work may be overlooked.

Despite the widespread use of the REF, it is not without its critics. By way of example, Sayer (2015) argued that panellists often review work outside of their area of expertise; in efforts to enhance their standings, universities may be selective in work submitted to panels and researchers may be pressured by their institutions to focus on work of likely appeal to assessors. Despite these and other possible limitations, quality assessments based on the number of times works are cited (h-index) correlate well with REF scores (Bruns & Stern, 2015; Clerides et al., 2011)—not that citations are the gold standard in accordance with which other measures should be evaluated. However, magnitudes of associations may vary by department. In some cases, they fail to attain statistical significance (Busso & Di Tollo, 2022). Overall, despite possible limitations, REF scores can be viewed as good measures of research quality. In fact, we believe that they should be the standard against which results of other forms of research assessment are compared.

In the current analysis, although the *CUG*, *The Guardian*, and *The Times* all produce league tables, the public data provided by *The Guardian* are used as a starting point. As noted above, this organization imputes values for missing data and provides the important variable, value added. Where appropriate, information provided by *The Guardian* is supplemented by material offered by the *CUG* and the Higher Education Statistics Agency. Given the previously articulated problem with reputational data, *The Times* is not included in the analyses.

In our analyses of institutions, we utilize data from 121 universities that are included in league tables for 2019, the last pre-pandemic year. Sometimes, however, complete data are unavailable for all variables. In such cases, analysis is based on a smaller number.

In this study, as for all league tables, the focus is on institutions, not individuals. This is an important distinction. Sometimes relationships established at the institutional level are not found when individuals are studied. The reverse is also true.

Institutional Level

Value added in learning. As noted above, *The Guardian*, in its league tables, provides a useful measure of value added in learning. *The Guardian* can use this metric as most students in the United Kingdom are admitted to university based on their "A-level" results (largely examination based), save for Scotland. Whatever the limitations of such forms of assessment, all students are evaluated in accordance with a common standard. As a result, differences between entry grades and type of earned degree may be viewed as a measure of value added by a particular university. A previous analysis demonstrated that grade inflation did not invalidate the use of this metric (Grayson & Grayson, 2023).

In Canada it is different. While curricula are developed primarily at the provincial level, assessment is mainly carried out in home schools. As a result, it is impossible to determine if, for example, an A grade in one school is equivalent to the same grade elsewhere.

Warrant for this scepticism is provided by several studies. In the province of Alberta, graduating high school students complete final examinations in their home schools. These are evaluated by their teachers. The students also write provincial examinations in the same subjects graded centrally. Performance on the latter is lower—for some subjects much lower—than on assessments of grades in home schools (Johnson, 2022). From one perspective, such differences call into question the validity of the home school assessments. Equally important, the University of Waterloo in the province of Ontario has shown that the association between school grades and university achievement varies drastically from one high school to the next (Cain, 2018). This finding can be taken as an indication of variation in high school standards. A further study, conducted at York University in Ontario, involving data from administrative records found little association between high school and university grades. In contrast to high school grades, measures of students' generic skills were a far better predictor of university accomplishment (Grayson, 2020).

In view of these limitations, in contrast to the A-levels in Britain, high school grades in Canada cannot be used as valid indicators of incoming students' expertise. Consequently, analyses of value added in learning as conducted by *The Guardian* are not possible. Nonetheless, were the data available, it would be reasonable to *hypothesize* that in Canada and elsewhere institutional factors that contribute to value added in learning would be comparable to those in the United Kingdom.

ANALYSIS

Institutional Level

Value added in learning. Consistent with the former qualifications and limitations, using U.K. data for 2019, we regressed value added in learning (VA) on student/staff ratio; excluding professorial and staff salaries, money spent on students (student services); entry grades; research quality as measured by the REF; and number of enrolled undergraduate students. Note that there was no institutional-level theory relevant to the selection of these variables. They were chosen heuristically. The results of their analysis are shown in Figure 1.

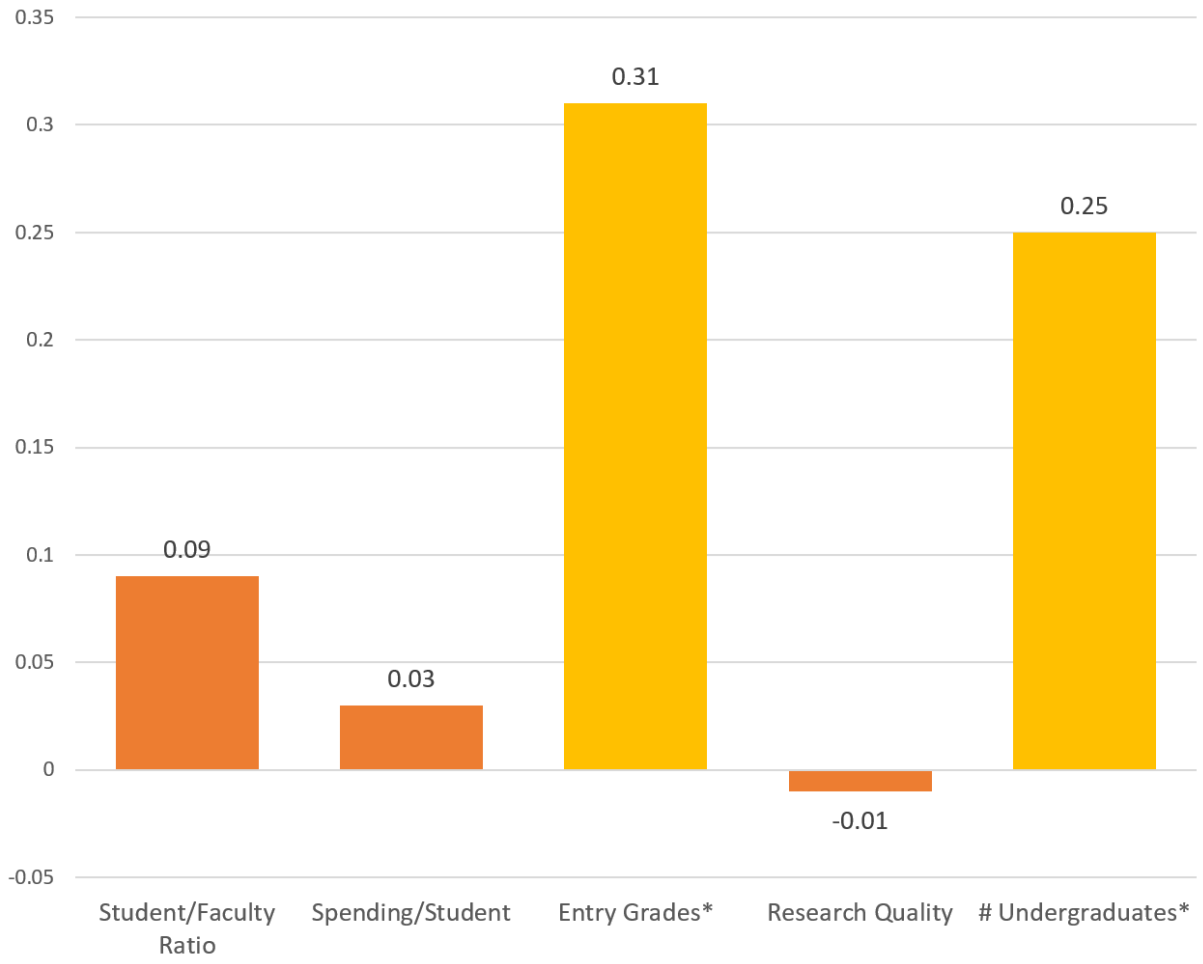
In the figure, bars represent beta coefficients. For the unfamiliar, the larger the beta, the greater the effect of the variable on the dependent variable—in this case, value added. It is important to note, however, that even large betas do not guarantee statistical significance.

In the figure, orange bars represent betas that are not statistically significant. Yellow bars show statistically significant betas. From the bottom left corner, we see that the overall regression model was statistically significant and that it explained 20% of the variance.

Most evident in the figure is that entering grades have the greatest positive effect on value added (beta = .31). Number of enrolled undergraduates is the second most important variable (beta = .25). It is important to note that the effect of the student/staff ratio is small (beta = .09) and not statistically significant. Similarly, expenditures per student (student services) have virtually no impact on value added (beta = .03). This measure also lacks statistical significance. The effect of the quality of faculty research is practically non-existent (beta = -.01) and does not achieve statistical significance.

The overall conclusion that can be derived from these data is that the best predictor of value added in learning is student quality at entry, particularly in large universities. This makes sense. The U.K. system has less classroom contact time and higher stakes assessments than Canada. Students with a track record of academic performance, as measured by A-level results, likely would be more able than others to make effective use of available learning resources. The ratio of students to staff, and the amount of spending on student services, is of no consequence.

Figure 1
Beta Values for VA



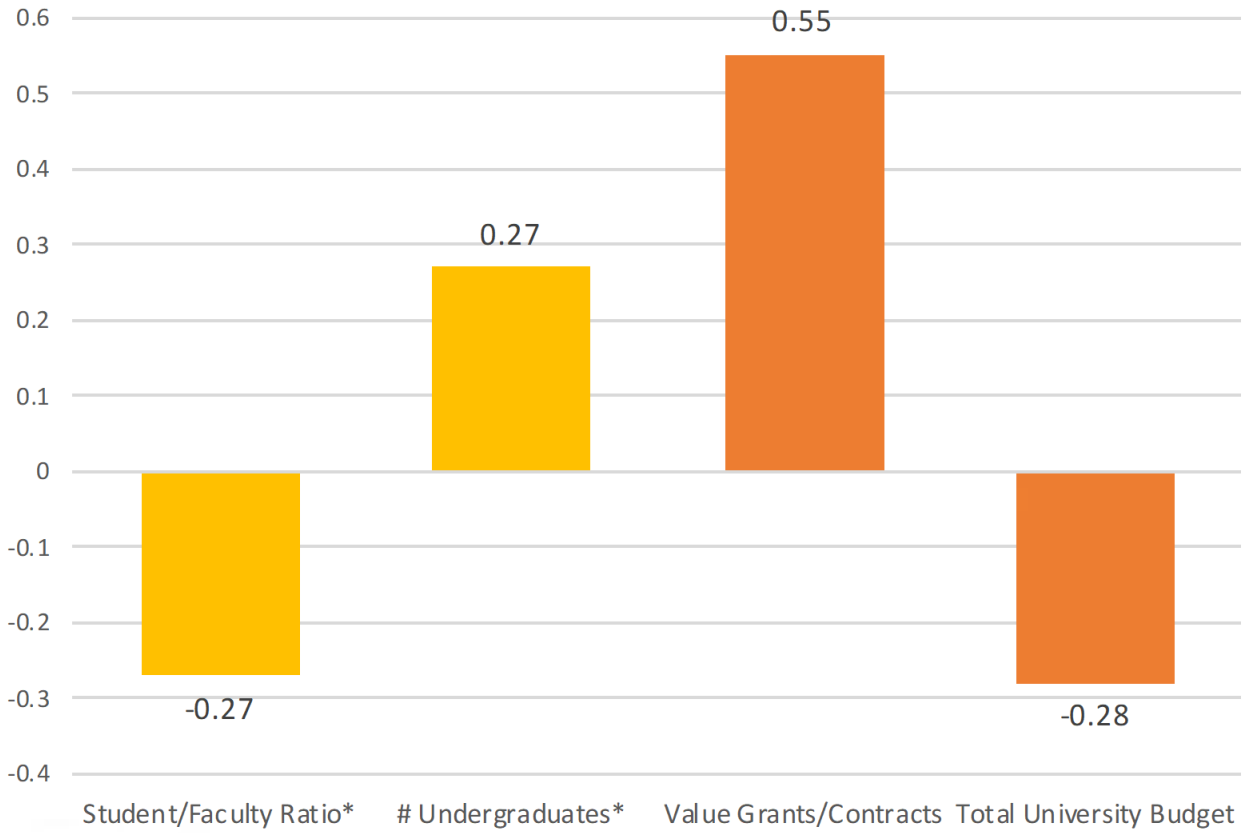
Note. R sq = .20; F p < .05; *t p < .05.
Yellow = Significant; Orange = Not significant.

This does not mean that the United Kingdom should take the lid off class size or stop spending on student services. Were such possibilities introduced, universities might reach levels after which value added would decline considerably. It also does not preclude the possibility that in any one university such variables would be of consequence for value added in learning.

As noted earlier, *Maclean's* bases some of its assessment of universities on students' receipt of academic awards and scholarships. However, based on U.K. data it would be unwarranted to assume that such distinctions point to the quality of the university. In view of foregoing findings, academic awards and scholarships could equally be a manifestation of entering student levels of achievement.

Research quality. In the absence of solid theory at the institutional level focusing on research quality, for heuristic reasons, we regressed REF scores (research quality) as described earlier on student/staff ratios, number of undergraduate students, total research grants and contracts, and total university budget. The results are shown in Figure 2.

Figure 2
Beta Values for Research Quality



Note. R sq = .43; F p < .05; *t p < .05.
Yellow = Significant; Orange = Not significant.
.

The most important finding in Figure 2 is despite a large beta, there is an absence of a statistically significant effect of total research grants and contracts (beta = .55) on research quality. The total university budget (beta = -.28) has no statistically significant relationship to the same. Only the student/staff ratio (beta = -.27) and number of undergraduates (beta = .27) affect research quality in a statistically significant way. In other words, all we can say with confidence is that relatively large universities and those with relatively low student/faculty ratios produce higher quality research than others.

This may seem counterintuitive. How can obtaining research support have no statistically significant relationship to research quality? The answer lies in the nature of the research process itself. Depending upon the discipline, excellent work can be completed with little or no research funding. Accordingly, grants and contracts may better reflect areas of faculty concentration than quality of research per se. For example, the costs of completing a national election study may be far greater than researching a book on the consequences of changes in immigration policy. Nonetheless, it would be hard to argue that the former would be of better quality than the latter, simply because it cost more.

The foregoing also contributes to a possible understanding of the finding that the lower the student/faculty ratio, the greater the research quality. Particularly in the absence of grants and contracts, which allow researchers to outsource certain research tasks, a lower student/staff ratio may reduce required teaching time. The time thereby saved might easily be committed to research.

Overall, via the REF, the U.K. system prioritizes research outputs. Research grants and contracts are inputs. The quality of research outputs should not be confused with their costs. Accordingly, based on the U.K. data, questions can be raised as to the validity of *Maclean's* and others' use of research grants as a possible proxy for research quality.

Discipline Level

Value added in learning. United Kingdom rankers make information readily available on their websites at both the university and departmental levels. This is very important. Although the point is not sufficiently emphasized by rankers, departmental levels of achievement as reflected in REF scores and value added may vary considerably from institutional scores found in league tables.

Space constraints preclude a complete discussion of all departments for which information is available. Instead, we focus on Departments of Politics and International Studies. Our intent in so doing is to provide an indication of ways in which data at the departmental level can be used. This discipline was chosen for analysis as its practitioners conduct both high- and low-cost research.

In Table 1, value added as reported by *The Guardian* is related to overall position in league tables. The table shows that there were many ties in value added. Therefore, the number of resulting value categories reduced to seven. Note that in this and following tables the number of universities exceeded the space in which to identify them; however, analysis is based on all institutions for which data were available.

Table 1
VA and Guardian *Rank*

Category	Institution	VA	*Guardian* rank
1	Aberdeen	9.0	20
	East Anglia	9.0	38
	Essex	9.0	85
	Goldsmiths'	9.0	105
	Lincoln	9.0	52
	LSE	9.0	5
	University College London	9.0	14
2	Bristol	8.0	11
	City	8.0	95
	Coventry University	8.0	26
	Hull	8.0	72
	King's College London	8.0	42
3	Aberystwyth University	7.0	49
	Bath	7.0	6
	Cambridge	7.0	3
	Edinburgh	7.0	13
	Kent	7.0	82
	Leeds Beckett	7.0	106
	Sheffield	7.0	31

Category	Institution	VA	*Guardian* rank
	SOAS	7.0	86
	Southampton	7.0	23
4	Birmingham	6.0	21
	Cardiff University	6.0	37
	Exeter	6.0	16
	Oxford	6.0	1
	Queen Mary	6.0	68
	Warwick	6.0	8
5	Durham	5.0	4
	Glasgow	5.0	12
	Keele	5.0	35
	Leeds	5.0	18
	Manchester	5.0	25
	Strathclyde	5.0	15
	Sussex	5.0	105
	Swansea University	5.0	24
	Westminster	5.0	117
	Winchester	5.0	94
	York	5.0	16
	Liverpool	4.0	39
6	Leicester	3.0	77
	Loughborough	3.0	7
	Nottingham	3.0	38
	Oxford Brookes	3.0	43
	Reading	3.0	62
7	Brunel University London	2.0	110
	Newcastle	2.0	51
	Queen's University	2.0	46
	Royal Holloway	2.0	32
	Stirling	2.0	28
	Mean VA	5.8	
	SD	2.2	
	r	0.0	
	p	> .05	

Table data show that there was variation between value added and overall *Guardian* ranks. For example, although Oxford rated number one overall, its Politics and International Studies department only placed in value added category 4. While the overall *Guardian* rank for Cambridge was 3, along with several other universities, its Politics and International Studies department was also in value added category 3. The overall correlation between value added and league table position, as seen below the table, was zero.

Clearly, any assumption that highly ranked U.K. universities offer students a better education in Politics and International Studies than others is not supportable. If anything, except for the London School of Economics (LSE), which has a *Guardian* score of 5 and a value added in the first category, there is little relationship between overall university placement in league tables and the value added of Politics and International Studies departments.

Perhaps more importantly, the correlation between research quality and value added in Politics departments was only .08 (not shown in table). Moreover, it was not statistically significant. This means that departments with high-quality research contribute no more than others to value added in learning.

The implication of the above for Canadians and others is as follows: Although *Maclean's* attaches importance to the number of student awards and scholarships as an implicit index of value added, in the light of U.K. data, this may be a questionable assumption. In addition, any notion that Departments of Politics and International Studies with high research quality impart more learning to students is unsustainable. As a result, we should assume null hypotheses with respect to these matters in examinations of Canadian and other universities. Unfortunately, in the absence of data comparable to that of the U.K., we are not able to test the resulting hypotheses.

Research quality. Although *The Guardian* ranks at least 121 institutions, only 56 of this number of departments elected to submit to REF 2021 under Unit of Assessment 19 (Politics and International Studies). This reflects institutional choices. In Table 2, quality of research (REF score) is related to the amount received by departments in grants and contracts from 2014 to 2020 and the overall rank assigned by *The Guardian*.

Table 2
Politics Research Quality, Funding, and Guardian Rank.

Category	Institution	REF	£ Research 2014-20 (000)	Guardian Rank
1	Strathclyde	3.6	10,111	15
	Royal Holloway	3.6	936	32
	LSE	3.5	18,521	5
	University College London	3.5	6,768	14
	Edinburgh	3.5	12,156	13
	Essex	3.5	3,161	85
	Queen Mary	3.5	3,606	68
2	Manchester	3.4	7,988	25
	Warwick	3.4	11,803	8
	Sheffield	3.4	7,290	31
	York	3.4	4,943	16
	Oxford	3.4	31,490	1
	Sussex	3.4	2,594	105

Category	Institution	REF	£ Research 2014-20 (000)	Guardian Rank
3	SOAS	3.3	2,421	86
	East Anglia	3.3	1,594	38
	Exeter	3.3	11,918	16
	Bristol	3.3	1,339	11
	Glasgow	3.3	3,807	12
	Reading	3.3	563	62
	King's College London	3.3	47,005	42
4	Cambridge	3.2	9,253	3
	Birmingham	3.2	7,396	21
	Cardiff University	3.2	1,999	37
	Nottingham	3.2	4,658	38
	St Andrews	3.2	2,836	2
	Aberystwyth University	3.2	2,780	49
5	Southampton	3.1	1,998	23
	City	3.1	2,028	95
	Durham	3.1	3,485	4
	Birkbeck College	3.1	718	
	Leeds	3.1	2,825	18
	Queen's University	3.1	5,026	46
	Kent	3.1	2,462	82
6	Liverpool	3.0	1,501	39
	Westminster	3.0	822	117
	Newcastle	3.0	1,136	51
	Leicester	3.0	1,930	77
	Brunel University London	3.0	705	110
	Bath	3.0	2,546	6
	Lincoln	3.0	691	52
7	Coventry University	2.9	7,048	26
	Swansea University	2.9	953	24
	Stirling	2.9	551	28
	Loughborough	2.9	665	7
	Goldsmiths'	2.9	661	105

Category	Institution	REF	£ Research 2014-20 (000)	Guardian Rank
8	The Open University	2.8	572	
	Oxford Brookes	2.8	420	43
	Keele	2.8	651	35
	Bradford	2.8	803	99
9	Kingston University	2.7	2,176	40
	Aberdeen	2.7	2,724	20
10	Leeds Beckett	2.6	135	106
11	Canterbury Christ Church University	2.5	158	114
	Winchester	2.5	0	94
12	Hull	2.3	483	72
13	Liverpool Hope	1.9	1	84
	Mean REF	3.08		
	SD	0.33		
	Mean £ Research	4,728K		
	SD	7,850K		
	r REF & £Research	0.39		
	p	< .05		
	r REF & Rank	-0.42		

The table shows that, because of ties, the REF scores of departments in the table reduce to 13 categories. It also reveals the amount received by departments in research grants and contracts as reported by HESA. The overall mean for grants and contracts was £4,728k (below table). Strathclyde, one of the two top departments, received £10,111k. Similarly placed Royal Holloway reported a much lower £936k. By contrast, Oxford, placed in the second REF category, received £31,490k. This sum is three and 34 times larger than that of Strathclyde and Royal Holloway, respectively! Similarly, Cambridge, in research quality category 4, reported £9,253k. This figure is slightly smaller than that of Strathclyde, but 10 times that of Royal Holloway. Despite their first-place position in quality of research in Politics and International Studies, Strathclyde and Royal Holloway have *Guardian* ranks of only 15 and 32, respectively.

Beneath the table we see that the correlation between politics REF scores and overall ranking for universities is a statistically significant -.42. The sign is negative because a high REF score places an institution toward the top of the overall placement, which begins with one. Importantly, although the correlation is statistically significant, it explains only 16% of the total variance. The reasons for this are the widely divergent scores for many universities in the table.

Also shown at the bottom of the table is the statistically significant correlation of .39 between REF position and research funding. Despite attaining statistical significance, the correlation only explains 15% of the variance.

The significance of these findings is that an overall U.K. university ranking may overshadow the research quality of individual departments. Part of this distortion may result from combining the REF results from departments to provide an

overall research score and the mixing of measures of quality with matters like student satisfaction, entering grades, and the value of research grants and contracts. Each contribute to the final league table ranking. Whatever the case, in Politics and International Studies departments, research funding and overall *Guardian* placement are poor reflections of research quality.

Unfortunately, neither *Maclean's* nor any other organization would be able to unmask the reality behind research grants and research quality in Canada and elsewhere. We simply lack a measure of research quality comparable to that of the United Kingdom. Were this not the case, we would have a better appreciation of research quality stripped of distorting measures, like the size of research grants. Unfortunately, in overall rankings, research quality could be overlooked because of low standing on less important matters included in the production of league tables.

In Sum

To clarify further, the figures in Tables 1 and 2 have been plotted in Figures 3, 4, and 5. In Figure 3, the stepped blue line represents, in descending order from left to right, the value added category of Politics and International Studies departments. Clearly, value added does not vary in a systematic way with university rank (the solid orange line).

In Figure 4, Politics and International Studies departments have been sorted in descending order of REF category from left to right. The stepped blue line demonstrates this distribution. Reflecting the preceding tables, associations among REF category, research funding (the orange wave), and *Guardian* ranking (grey bars) are erratic and weak.

In Figure 5, institutions are sorted in descending value of REF category from left to right. As above, the stepped blue line represents this distribution. Data summarized in the chart indicate that there is virtually no connection between REF and value added (solid orange line) categories.

Figure 3
Politics: Relationship Between Politics VA Category and Guardian University Rank

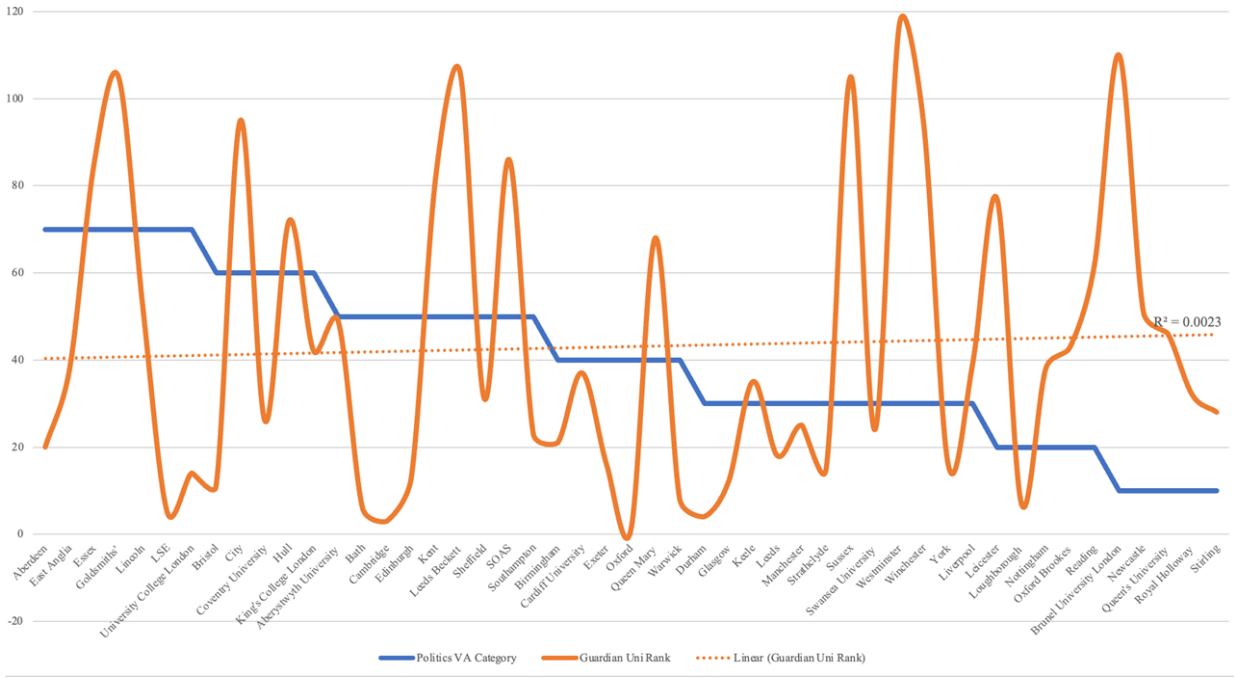


Figure 4
Politics: Relationship Among REF Category, Funding, and Guardian Rank

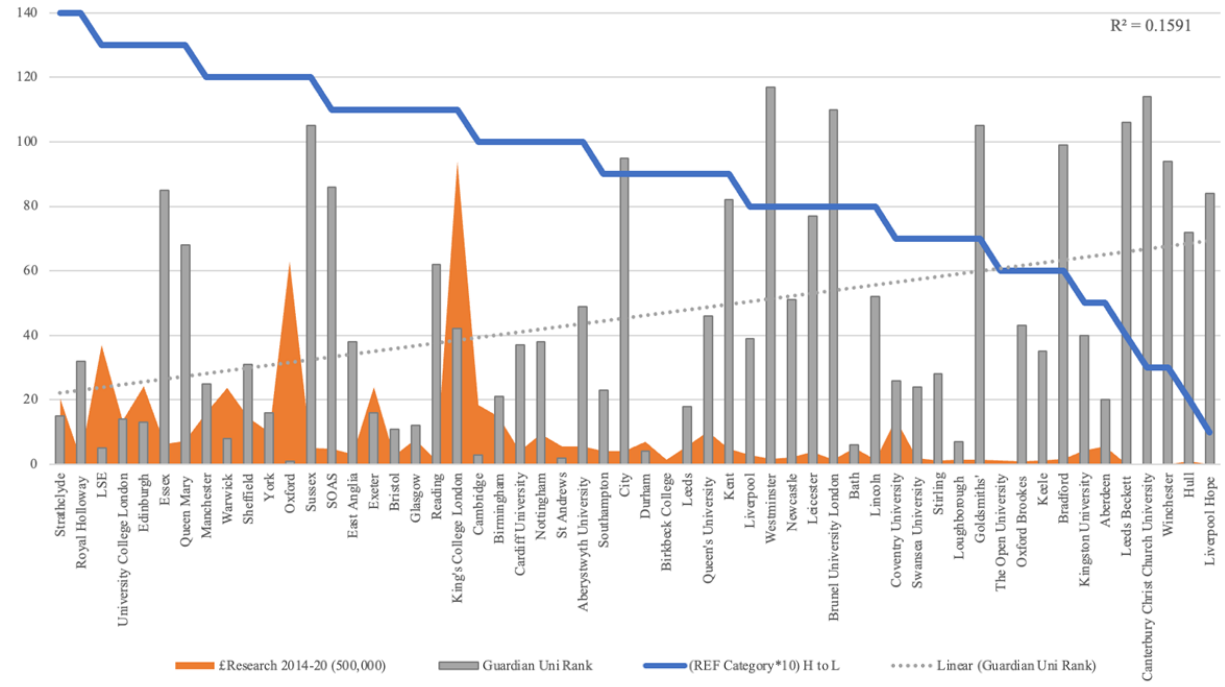
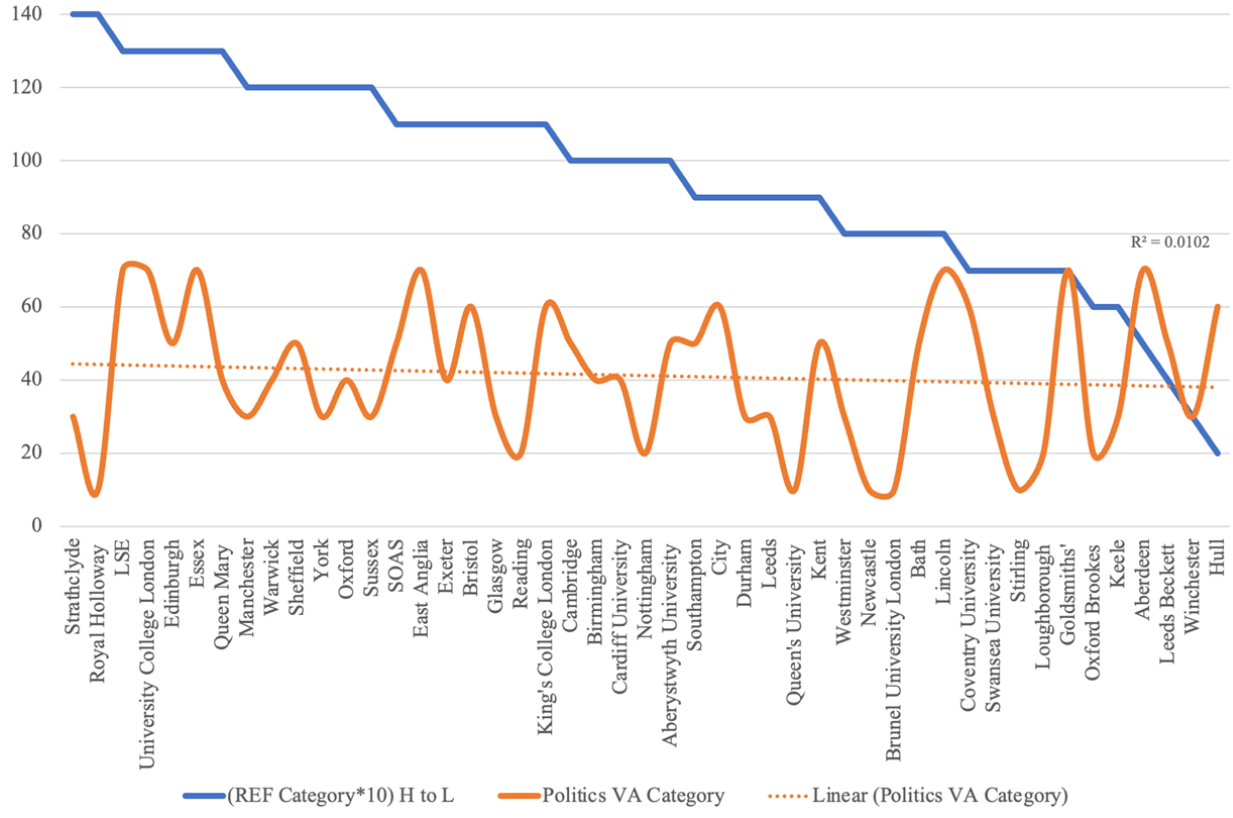


Figure 5
Politics: Relationship Between Politics REF Category and VA Category



CONCLUSION

The findings presented here have direct implications for how universities, in Canada and elsewhere, are evaluated and how rankings are used by policy makers, institutional leaders, and the public. Drawing on rigorous U.K. data, this article exposed critical weaknesses in common league table indicators. Limitations were particularly evident when research funding and student awards were used as proxies for institutional quality.

In the United Kingdom, at both the institutional and one departmental level, we found little consistent relationship between funding and actual research excellence, and institutional rank and value added in learning. These results suggest that in many countries in which university rankings are undertaken, metrics used in ranking systems may distort public understanding and institutional priorities. Regarding the latter, reliance on easily quantifiable inputs over validated outcomes risks incentivizing performance on superficial measures, rather than on substantive educational or scholarly contributions.

In light of these facts, in Canada and other countries, governments, ranking agencies, and universities should adopt a more transparent, evidence-based approach to assessing institutional performance. Where possible, measures of student learning and rigorous peer-reviewed research assessment—like the United Kingdom's REF and value added metrics—should serve as models for more meaningful evaluation. In their absence, stakeholders must treat current rankings with caution. They are unable to capture the complex realities of university quality.

CONFLICT OF INTEREST

The corresponding author states that there is no conflict of interest.

REFERENCES

Amsler, S. S., & Bolsmann, C. (2012). University ranking as social exclusion. *British Journal of Sociology of Education*, *33*(2), 283–301. https://doi.org/10.1080/01425692.2011.649835

Boliver, V. (2015). Are there distinctive clusters of higher and lower status universities in the UK? *Oxford Review of Education*, *41*(5), 608–627. https://doi.org/10.1080/03054985.2015.1082905

Bowden, R. (2000). Fantasy higher education: University and college league tables. *Quality in Higher Education*, *6*(1), 41–60. https://doi.org/10.1080/13538320050001063

Brankovic, J. (2021, March 22). The absurdity of university rankings. *LSE Impact Blog.* https://blogs.lse.ac.uk/impactofsocialsciences/2021/03/22/the-absurdity-of-university-rankings

Broecke, S. (2015). University rankings: Do they matter in the UK? *Education Economics*, *23*(2), 137–161. https://doi.org/10.1080/09645292.2012.729328

Bruns, S., & Stern, D. (2015). *Research assessment using early citation information*. Crawford School Research Paper No. 15-0. http://dx.doi.org/10.2139/ssrn.2602151

Busso, A., & di Tollo, G. (2022). Prediction of UK research excellence framework assessment by the departmental h-index. *European Journal of Operational Research*, *296*(3), 1036–1049. https://doi.org/10.1016/j.ejor.2021.05.006

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

Calderon, A. (2021, March 8). *Lesson from the subject rankings is that funding is key*. University World News. https://www.universityworldnews.com/post.php?story=20210308123312913

Chapman, K. (1996). Entry qualifications, degree results and value-added in UK universities. *Oxford Review of Education*, *22*(3), 251–264. http://www.jstor.org/stable/1050585

Chevalier, A., & Jia, X. (2015). *Subject specific league tables and students' application decisions.* Institute for the Study of Labour. https://www.iza.org/publications/dp/8869

Clerides, S., Pashardes, P., & Polycarpon, A. (2011). Peer review vs. metric-based assessment: Testing bias in the RAE ratings of UK economics departments. *Economica*, *78*(311), 565–583. https://doi.org/10.1111/j.1468-0335.2009.00837.x

Complete University Guide. (2021). *Your choice, your future*. https://www.thecompleteuniversityguide.co.uk

David, M. (2016). Fabricated world class: Global university league tables, status differentiation and myths of global competition. *British Journal of Sociology of Education*, *37*(1), 169–189. https://doi.org/10.1080/01425692.2015.1096190

Dill, D. D., & Beerkens, M. (2005). Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. *Higher Education*, *49*, 495–533. https://doi.org/10.1007/s10734-004-1746-8

Dolton, P., & Makepeace, G. (1982). University typology: A contemporary analysis. *Higher Education*, *14*(3), 33–47.

Drewes, T., & Michael, C. (2006). How do students choose a university? An analysis of applications to universities in Ontario, Canada. *Research in Higher Education*, *47*(7), 781–800. https://doi.org/10.1007/s11162-006-9015-6

Easley, J., Strawderman, L., Babski-Reeves, K., Bullington, S., & Smith, B. (2021). Perceived quality factors in higher education. *Quality in Higher Education*, *27*(3), 306–323. https://doi.org/10.1080/13538322.2021.1909210

Foley, B., & Goldstein, H. (2012). *Measuring success: League tables in the public sector*. The British Academy. https://www.thebritishacademy.ac.uk/publications/measuring-success-league-tables-public-sector/

Gibbons, S., Neumayer, E., & Perkins, R. (2015). Student satisfaction, league tables and university applications: Evidence from Britain. *Economics of Education Review*, *48*, 148–164. https://doi.org/10.1016/j.econedurev.2015.07.002

Grayson, J. P. (2020). The Emperor's New Clothes: Maclean's, NSSE, and the inappropriate ranking of Canadian universities. *Canadian Journal of Higher Education*, *50*(3), 14–35. https://doi.org/10.47678/cjhe.vi0.188713

Grayson, K., & Grayson, J. P. (2023). University quality, British league tables and student stakeholders. *Quality in Higher Education*, *30*(2), 257–284. https://doi.org/10.1080/13538322.2023.2235732

Gruber, T. (2014). Academic sell-out: How an obsession with metrics and rankings is damaging academia. *Journal of Marketing for Higher Education*, *24*(2), 165–177. http://dx.doi.org/10.1080/08841241.2014.970248

Harvey, L. (2008). Rankings of higher education institutions: A critical review. *Quality in Higher Education*, *14*(3), 187–207. https://doi.org/10.1080/13538320802507711

Hiely-Rayner, M. (2016, May 23). Methodology behind the Guardian University Guide 2019. *The Guardian*. https://www.theguardian.com/education/2016/may/23/methodology-behind-the-guardian-university-guide-2017

Higher Education Statistics Agency. (2021). *Higher Education Statistics Agency*. https://www.hesa.ac.uk

Jarocka, M. (2012). University ranking systems – from league tables to homogeneous groups of universities. *International Journal of Educational and Pedagogical Sciences*, *6*(6), 1377–1382.

Johnson, D. R. (2022, April 28–29). *Measuring grading standards at high schools: A methodological contribution with an example* [Conference presentation]. National Bureau of Economic Research Education Meeting, Stanford, CA, United States.

Jung, C. S., Toutkoushian, R. K., & Teichler, U. (Eds.). (2011). *University rankings: Theoretical basis, methodology and impacts on global higher education*. Springer.

Lynch, K. (2015). Control by numbers: New managerialism and ranking in higher education. *Critical Studies in Education*, *56*(2), 190–207. https://doi.org/10.1080/17508487.2014.949811

Maclean's. (2023). *Maclean's university rankings*. https://education.macleans.ca/rankings

Marginson, S. (2006). Dynamics of national and global competition in higher education. *Higher Education*, *52*, 1–39. https://doi.org/10.1007/s10734-004-7649-x

Marginson, S. (2014). University rankings and social science. *European Journal of Education*, *49*(1), 45–59. https://doi.org/10.1111/ejed.12061

Mayhew, M. J., Rockenbach, A. N., Bowman, N. A., Seifert, T. A., Wolniak, G. C., Pascarella, E. T., & Terenzini, P. T. (2016). *How college affects students: 21st century evidence that higher education works* (Vol. 3). Jossey-Bass.

McManus, R., Haddock-Fraser, J., & Rands, P. (2017). A methodology to understand student choice of higher education institutions: The case of the United Kingdom. *Journal of Higher Education Policy and Management*, *39*(4), 390–405. https://doi.org/10.1080/1360080X.2017.1330806

Moed, H. F. (2017). A critical comparative analysis of five world university rankings. *Scientometrics*, *110*(2), 967–990. https://doi.org/10.1007/s11192-016-2212-y

National Student Survey. (2020). *The National Student Survey: Consistency, controversy and change*. Office for Students. https://www.officeforstudents.org.uk/publications/the-national-student-survey

National Student Survey. (2021). *National Student Survey 2021*. https://www.thestudentsurvey.com

Ordorika, I., & Lloyd, M. (2014). International rankings and the contest for university hegemony. *Journal of Education Policy*, *30*(3), 385–405. https://doi.org/10.1080/02680939.2014.979247

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: Findings and insights from twenty years of research*. Jossey-Bass.

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research* (Vol. 2). Jossey-Bass.

Research Excellence Framework. (2014). *Assessment criteria and level definitions*. https://2014.ref.ac.uk/panels/assessmentcriteriaandleveldefinitions/index.html

Research Excellence Framework. (2021). *Research Excellence Framework 2021*. https://www.ref.ac.uk

Sayer, D. (2015). *Rank hypocrisies: The insult of REF*. SAGE.

Schleicher, A. (2015). *Value-added: How do you measure whether universities are delivering for their students? HEPI 2015 annual lecture*. Higher Education Policy Institute. https://www.hepi.ac.uk/wp-content/uploads/2016/01/Andreas-Schleicher-lecture1.pdf

Smith, G. F. (1993). The meaning of quality. *Total Quality Management*, *4*(3), 235–244. https://doi.org/10.1080/09544129300000038

Soysal, Y. N., Baltaru, R., & Cebolla-Boado, H. (2020). *Meritocracy or reputation? How university rankings matter.* OSF Preprints. https://osf.io/preprints/socarxiv/f6by3

Zilvinskis, J., & Rocconi, L. (2018). Revisiting the relationship between institutional rank and student engagement. *The Review of Higher Education*, *41*(2), 253–280. https://dx.doi.org/10.1353/rhe.2018.0003