

Article

Enhancing the Positive Impact Rating: A New Business School Rating in Support of a Sustainable Future

Kathleen Rodenburg , Taimoor Rizwan, Ruifeng Liu and Julia Christensen Hughes

Gordon S. Lang Business School of Business and Economics, University of Guelph,
Guelph, ON N1G 2W1, Canada; trizwan@uoguelph.ca (T.R.); ruifeng@uoguelph.ca (R.L.);
jchriste@uoguelph.ca (J.C.H.)

* Correspondence: krodenbu@uoguelph.ca

Abstract: Business school rankings are “big business”, influencing donors and potential students alike, holding much sway over decanal and faculty priorities, particularly with respect to the curriculum as well as the focus and destination of research publications (i.e., in so-called “top” journals). Over the past several years, the perverse effects of these priorities have begun to be acknowledged, and new ratings and ranking systems have emerged. One promising newcomer is the Positive Impact Rating (PIR), which uniquely and exclusively focuses on student perceptions of their business school’s priorities and the learning experience. In addition, it organizes schools by tier, in an effort to foster collaboration and continuous improvement, as opposed to ranked competition. If this new approach is to achieve its stated objective and help shift the focus of business schools to developing future business leaders and research output in alignment with a more sustainable world (and the United Nations Sustainable Development Goals), it is essential that the metrics used be and be perceived as both valid and reliable. The current research aims to make a contribution in this regard, analyzing the results at one business school in detail and making recommendations for strengthening these aims. Results show that the parametric properties of the survey are highly interrelated, suggesting that the predictive utility of the separate elements within the scale could be improved. Additionally, biases in scores may exist depending on where the responses are collected and who solicited them, as well as the students’ perception of their overall academic experience and on socio-cultural factors.

Keywords: Positive Impact Rating (PIR); Sustainable Development Goals (SDGs); rankings; ratings; biases



Citation: Rodenburg, K.; Rizwan, T.; Liu, R.; Christensen Hughes, J. Enhancing the Positive Impact Rating: A New Business School Rating in Support of a Sustainable Future. *Sustainability* **2021**, *13*, 6519. <https://doi.org/10.3390/su13126519>

Academic Editor: Donato Morea

Received: 6 May 2021

Accepted: 4 June 2021

Published: 8 June 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Much has been written and debated about business school rankings over the past several years, acknowledging their limitations and offering suggestions for improvement. In particular, the metrics used by traditional rankings have been found wanting. Perverse effects on faculty and decanal priorities have been identified, including incenting behaviors that are at odds with the achievement of the United Nations Sustainable Development Goals (SDGs) [1,2] and the 2030 agenda. In response to this recognition, a new rating system was launched in 2019 known as the Positive Impact Rating (PIR) [3]. The PIR uniquely centers on students’ perceptions of their business school’s governance and culture, as well as the extent to which programs and learning methods have prepared them to pursue careers of purpose. The outcomes have also been uniquely tallied and presented, with business schools arranged into one of five tiers (from beginning to pioneering) as opposed to being ranked one against the other. This rating approach was intended to help foster collaboration and continuous improvement among business schools as opposed to competition.

According to its founders, the purpose of the PIR is to “help speed up the transformation” towards orienting “teaching, research and outreach activities towards social impact

and sustainability” [3] (p. 6). In order for the PIR to achieve its goals, including becoming broadly perceived as a reliable and valid assessment instrument, it is essential that its metrics and approach be held to a high standard, and further, that its various dimensions (currently energizing, educating, and engaging) are statistically supported.

This paper reports on the results of a study conducted at one business school, the Gordon S. Lang School of Business and Economics at the University of Guelph, Canada, during the 2020/21 academic year. The PIR scale as a benchmark in theory realizes the importance of business schools and their role in achieving the 17 SDGs [4] and therefore was selected as an aligned measurement tool for Lang given its Principles of Responsible Management Education (PRME) Champion School status. A PRME champion school is recognized by the United Nations Global Compact as leaders with a commitment to responsible business education. The scale was designed to assess student perceptions of their business school’s positive impact on society across several factors, including school governance, culture, programs, learning methods, student support, the institution as a role model, and its level of public engagement [3]. As a potential measurement tool that holds promise to overcome the perverse effects created by traditional business school ranking systems, it is essential that the metrics used be—and be perceived as—both reliable and valid. To this end, this study analyzes the efficacy of the PIR scale empirically to test the following null hypotheses:

Hypothesis 1 (H1). *The PIR is a reliable instrument as estimated by the coefficient alpha (Cronbach’s alpha).*

Hypothesis 2 (H2). *The confirmatory factor analysis supports the PIR as a valid instrument. A relationship exists between the observed responses and their underlying latent constructs; specifically, the responses to the survey support the model construct (Energizing, Engaging, Educating).*

Hypothesis 3 (H3). *There is a selection bias in channeling the data collection for the PIR through student organizations engaged in the sustainability field.*

Hypothesis 4 (H4). *There is a selection bias in collecting PIR data from students in courses linked to sustainability.*

Hypothesis 5 (H5). *Demographics or socio-cultural characteristics of the student influence PIR responses.*

Our paper begins in Section 2 with a brief synopsis of two events, which focused on the shortcomings of traditional business school rankings, held in Davos, during the 2019 World Economic Forum (WEF). In Section 3, the observations made at WEF are compared with support from the extant literature. Next in Section 4, we summarize our methods and results. Our discussion in Section 5 centers on the need to ensure that the measurement criteria for a chosen rating system support goals that enable a business school to achieve its vision and mission. In doing so, when used effectively as a recruitment tool, they could help attract students whose goals and values are aligned, strengthening organizational culture and further supporting the school’s ability to achieve its objectives. It is suggested that recruiters of recent graduates understand explicitly the criteria used to construct rankings and recognize that a graduate from a ‘top-ranked business school’ may not have the qualities assumed. In the spirit of future enhancements to this promising scale, we made the following recommendations: (1) a re-evaluation of the survey questions that informed the three separate elements (energizing, educating, and engaging) to ensure that these questions are measuring three distinct and separate aspects of a school’s positive societal impact; (2) a deliberate and broad-based distribution method for the survey to ensure participation by both highly engaged and less engaged students; and (3) an additional case study at a different school whose mission does not include the achievement of the 17 SDGs in order to compare results based on contrasting goals and student demographics

to further confirm its reliability. Finally, we discuss the contributions to the literature and the efficacy of the PIR scale, the implications for business practitioners and business school leaders, the limitations to the current study, and future research enhancements.

2. Challenges Levelled at Traditional Business School Rankings

During the 2019 WEF, an event on business school rankings was convened at the invitation of Corporate Knights and the UN Global Compact (and supported by the Globally Responsible Leadership Initiative (GRLI) [5]). Here, business school deans of PRME alongside deans and business leaders recognized by Corporate Knights, as globally leading sustainability champions, considered the need for significant change in business school rankings.

The event began with a presentation by the authors of the report ‘Business School Rankings for the 21st Century’ [2]. Based on an extensive literature review and focus groups with key stakeholders, Pitt-Watson and Quigley [2] suggested that business schools exert tremendous influence on society (through the knowledge and actions of their graduates). The priorities and actions of business schools, in turn, “appear to be greatly influenced by business school rankings” [2] (p. 23). Yet, the metrics of business school rankings do not appear to be in alignment with the needs of society, including the development of “a sustainable, inclusive 21st century economy” [2] (p. 23). More specifically, Pitt-Watson and Quigley concluded that the metrics used by traditional business school rankings fall short across several domains, including [2] (p. 23) “(a) salary overemphasis; (b) business schools penalized in the rankings for turning out graduates who work for non-profits; (c) course content not evaluated; and (d) teaching quality, sustainability and business ethics minimized or absent.” They concluded with a call for the re-imagination of rankings to “encourage these institutions to educate managers equipped to address the challenges of this era” [2] (p. 23).

Next, was a presentation by Katrin Muff on the newly developed Positive Impact Rating (PIR), scheduled for release in 2020. Muff explained how the design of the PIR was intended to respond to the Pitt-Watson and Quigley [2] critique, with its focus on the perceptions of current students, with respect to “the quality and nature of their learning experience and the culture of the business schools in which they study” [5] (p. 1). Schools in attendance were encouraged to participate in the PIR and join the celebration of its release during the World Economic Forum in 2020.

Following this, discussion groups considered three questions: Why do rankings matter and to whom? What is most unfortunate about current rankings? And what is the desired future state? A synthesis of the perceptions of the participants suggested that:

“[R]anking systems (with notable exceptions—such as Corporate Knights) have had perverse (unintended) consequences on the focus of faculty research, curricular and pedagogical innovation, and the student experience (particularly for undergraduate students). Driven by the desire to be well-ranked (with the concomitant rewards that such rankings engender—such as significantly enhanced brand and credibility amongst potential donors, faculty, students and senior university administrators), business schools have been strongly incited to “play the game” and engineer results, particularly in the areas of student salaries and faculty research.” [5] (p. 1)

Other observations included that business school rankings inordinately focus on MBA programs, which can deprive large undergraduate programs of needed attention and resources. Additionally, publication lists, such as the fifty journals included in the Financial Times ranking (the FT50), can influence who gets hired, as well as promotion and tenure decisions. The problem with the latter was underscored by Dyllick [6], who reported that journal rankings such as the FT50 contain considerable bias, privileging English speakers from Europe and North America who provide disciplinary-based explanations of past development, as opposed to addressing pressing societal issues, including in interdisciplinary ways.

During the 2020 WEF, a second Deans Multi-Stakeholder Dialogue at Davos took place. That year's event featured the launch of the PIR (with deans from top-rated schools in attendance), a panel discussion with a representative from Corporate Knights, the PIR, and the Financial Times. These group discussions centered around three key topics: (1) participant reaction to the PIR; (2) perceptions of changes to other rankings that are underway; and (3) creating a wish-list for further change [7].

Findings from the discussions suggested that there was broad support for the PIR and its focus on student perceptions, as well as its five rating bands (beginning, emerging, progressing, transforming, and pioneering schools). Participants supported the potential for this approach to help foster collaboration amongst the rated schools. Some concern was also expressed about the potential replicability of the results, given a relatively low bar for response rates (i.e., a minimum of 30 responses) and the method by which the survey was promoted to potential participants (via Oikos International, Net Impact, and local student leaders). The suggestion was made that future iterations should endeavor to ensure a demographically diverse group of students, from multiple programs and year levels, from those in leadership positions and otherwise, to enhance the reliability of the results.

Observations and recommendations for improving traditional rankings included making them "more equitable and inclusive", "embracing continuous improvement", "valuing teaching and learning", valuing "emerging inter-disciplinary journals and more accessible forms for research and dissemination", and developing "mechanisms for reporting on contributions to the UN's 2030 agenda" [8] (p. 3).

3. The Extant Literature

While much of the focus at the events in Davos was on the perceptions and lived experiences of participants, the report by Pitt-Watson and Quigley [2] contained an extensive literature review that helped inform these perceptions. The authors established that two primary methods are used for the evaluation of business schools; accreditation agencies (ex: AACSB, EQUIS) and media organizations (the Financial Times, the Economist, etc.), collectively referred to as "ranking publications". While Accreditation bodies are focused on improving business education, "ranking publications", on the other hand, are focused on benchmarking business schools. The literature is divided into three subsections. The first body of literature provides compelling reasons why a business school would be motivated to participate in rankings. The second section highlights the unintended consequences of pursuing rankings. In the final subsection, recent literature points to the need for a better understanding of the criteria used to construct traditional rankings and the need for change.

3.1. Motivation for Business School Participation in Rankings

Published ranking systems originated from the need for business schools to be more forward-facing toward their customers (students and businesses). According to Khurana [9], they worked. Business school rankings, such as the FT50, Times Higher Ed (THE), QS World University (QS), and U.S. News have had considerable influence over student choice. Research has shown that improvements in the ranking of a business school led to a surge in the number of applicants in the coming years [10] and that these rankings were more influential than other types of media in helping potential MBA applicants determine their preferred school [11]. Elebeck [12] found that students who graduated from highly ranked schools performed better and had higher salaries upon graduation [13].

Another reason for rankings being increasingly valued was due to the internationalization of management education and the fact that business programs were now being marketed to potential students around the world [14]. Additionally, recruiters used these rankings to target potential employees [10]. In other research, rankings have been shown to drive resources from external partners when they too value the hierarchy of prestige in higher education [15]. However, as Pitt-Watson and Quigley [2] and others have pointed out, the metrics being used to come up with these rankings may be troublesome.

3.2. The Unintended Consequences of Published Rankings

Many published ranking systems measure student salaries and progression after graduation, recruiter opinions, placement success, and in some cases intellectual capital as defined by research publications [10]. For example, the Financial Times focuses 50% of its ranking criteria on the weighted salaries of alumni as well as the publication record of faculty members in selected academic journals (FT50). Very few measured how effective the school is at teaching intended content and the required skills necessary to be successful business leaders [2]. In most cases, the metrics that are used in determining the rankings are generally not well known and can differ markedly in what they emphasize. It is important to know the criteria, as research has shown that ratings and rankings influenced both school and student behavior and how faculty and administrators assess the reputation and prestige of the institution [15–19].

Given the financial implications that accompany the top ranks, Athavale et al. [20] noted that some deans were at risk of losing their jobs if their published rankings fell. As a result, these measurement systems had the power to influence organizational mission, strategy, personnel and recruitment decisions, and public relation priorities [21,22].

Key findings suggested that ranking methodologies, i.e., the use of subjective weights and attributes, were problematic and were open to gaming [23]. For example, a school that valued rankings based on publication in certain journals puts pressure on faculty to publish in these designated 'A' journals [24]. The standard favored by 'A' category journals devalued the research published elsewhere, regardless of the content and its contribution [24]. Failure to demonstrate publishing power in these supposedly 'elite' journals risked being put on probation, not being eligible for promotion, or being given a contract for teaching only [24]. Rynes [25] wrote that the prestigious rankings market themselves as measuring research productivity, but in essence, they only measure the publications in high-impact-factor journals that include citations of only a constrained number of journals that the systems recognize. Rather than creating new and innovative research, the rankings seemed to be focused on categorizing academics into those who are successful in garnering academic prestige and those who are not [26]. One impact was that innovation in research was discouraged. Priorities have focused on improved rankings at the expense of furthering knowledge [27–29].

Alder and Harzing [26] found that the pressure to publish in top journals has also turned senior faculty's attention away from mentoring and coaching young colleagues toward publishing their own research. Rankings have been found to influence hiring and promotion decisions [17,18]. Bennis and O'Toole [30] noted that rankings are so powerful that even "rational, well-informed, well-intentioned" faculty and institutional administrators would take and support actions that focused on the interest of achieving high rankings, even when such actions undermined strategy, teaching objectives, service goals, and, consequently, society.

More recent contributions have further extended the understanding of the perverse effects of traditional rankings and their metrics [23,31,32]. For example, Dearden, Grewal, and Lilien [31] highlighted that ranking publications, in addition to offering objective information, also affect the prestige of these schools, which in many cases acts against the preferences of students. They concluded that the research capabilities of business schools, which form a heavy component of ranking metrics, introduce the risk of problematic research practices. Hall and Martin [32] also found that pressure to publish in leading journals was associated with higher rates of academic misconduct. Drawing on other studies, they were able to find examples of "blatant misconduct", "questionable conduct", and "inappropriate conduct", given the pressure to publish [32].

Johnson and Orr [33] conducted interviews with 70 professors, researchers, and external stakeholders to better understand what meaningful business research meant for them and whether current research was viewed as "impactful". The results showed that opinions varied and were described by researchers and business practitioners as either "a dilution of scholarly rigor and academic autonomy; as a tokenistic effort at practitioner

engagement; or as a welcomed development that enables scholars to embrace the pursuit of actionable knowledge" [33] (p. 569). Some business practitioners viewed academic research as focusing on long-term strategic outcomes, and therefore not applicable given the fast-paced, constantly changing aspects of today's business environment. One business practitioner shared that they were "staggered at the trivial nature of papers published and the almost invisible veneer of knowledge that seems to be added to the world as we know it" [33] (p. 566). Some identified initiatives such as the Research Excellence Framework (a research impact evaluation of British higher education institutions) prevented faculty from doing practical research as it was classified as less valuable. A common answer by business leaders in other studies was that they did not consider the research in question as relevant [34,35].

3.3. *Negating Unintended Consequences of Rankings*

Several universities and business schools have refused to take part and have requested not to be included in the rankings [36]. However, research has shown that non-participation could lead to organizational illegitimacy [37]. Hence many business schools, although in disagreement with the methodologies and the consequences of pursuing these rankings, still participate [37], shifting their priorities to align. Unfortunately, this shift has made business schools less adaptive to current demands by students and other stakeholders. Cutter [38] found that for the past couple of years, MBA applications have seen a sharp decline in the U.S. [39]. Technological advancements have enabled customized learning programs more aligned with student needs, but unfortunately, business schools have not been leading this change [38]. Furthermore, Cutter [38] noted that business schools tend to be viewed internally by administration as focusing on "big money" initiatives and are less serious about solving problems and advancing knowledge for the discipline. So, what are we left with? A marketing tool that is essentially driving business schools out of business.

As participation in rankings is largely a voluntary undertaking, careful consideration is needed when selecting any particular system; ideally, it should be one that supports the vision, objectives, and core values of the institution. The specific metrics used should motivate administrators, faculty, and students to take actions aligned with the strategic goals and drive internal and external resources in this direction. Complicating the ranking system selection process are the imperfections and potential biases within each scale. There is an overwhelming lack of transparent information on both the validity and accuracy of these measures, which can have tremendous ethical and economic consequences. For example, Morgeson and Nahrgang [40] found that BusinessWeek's rankings placed little weight on student learning outcomes and/or the benefits to society, with most of their emphasis on "economic returns from the education" [39] (p. 31).

Business schools have a major influence on society [41]. Although some have recognized their potential role in achieving the UN SDGs, many have not been engaged [42]. Edwards et al. [43] highlighted that even though accreditation bodies have started putting increasing emphasis on sustainability learning outcomes, sustainability learning is complex and requires an interdisciplinary approach. That said, ranking publications have made some efforts in this area [44]. For example, the Financial Times recently added a weight for corporate social responsibility (CSR) in its ranking criteria. Unfortunately, the weighting they assigned for courses that teach CSR amounted to just 3% of the total criteria. To evoke meaningful change, there is a need for a ranking system that measures the degree of integration of this critical content across all courses offered and research activities [45]. The inclusion of these metrics is important; studies have shown that if a sustainability-focused curriculum is implemented in an effective manner, it can raise awareness and change the behavior of students [46].

The Corporate Knights and American Association for Sustainability in Higher Education (ASHE) have introduced ranking systems that are in stark contrast to the aforementioned scales. Specifically, they take into account initiatives that target a broad group of stakeholders [45]. Corporate Knights [47] aligns its criteria to measure a school's contri-

bution to the advancement of the 17 SDGs by identifying the integration of sustainability and ethics within its initiatives, curriculum, and academic research. The AHSE ranking system, called STARS (The Sustainability Tracking, Assessment, and Rating System), is a transparent, self-reporting framework for colleges and universities to measure their own sustainability performance.

A new and promising entrant to the “published ranking” forum is the Positive Impact Rating (PIR) scale. The PIR was developed in response to the need to integrate both ethics and sustainability into business schools’ curriculum and initiatives [3]. It is defined as a “rating scale”, not as a ranking, as it does not pit schools against each other, but rather provides a tool to measure how successful a school has been at “educating, energizing, engaging” [3] (p. 9) students in topics, concepts, and initiatives that focus on “business as a force for good”, recognizing the role business plays in achieving the 17 SDGs. The PIR survey is completed by students within the institution to measure the social impact and the quality of sustainability initiatives from their perspective. In addition to providing a benchmark for schools focused on this area, the rating can be used by prospective students when deciding on which business school they wish to attend. “Many students care deeply about making a positive difference through their professional lives, yet they do not necessarily know the right business school to get prepared” [3] (pp. 7–8). Not only does it turn the focus toward teaching effectiveness, but it also helps attract students who are aligned with the school’s mission and vision, enhancing the organizational culture. Aligning student values with a business school’s goals serves to increase the probability of the institution achieving its objectives. Although a major shift in how we measure success, inherent in all scales is the possibility of biases that may lead to unintentional results.

4. Methodology

To collect data for this study, undergraduate and graduate business students from a public Canadian university (Gordon S. Lang School of Business and Economics) completed a questionnaire to test for possible selection biases resulting from the way the PIR rating responses were collected and to see how these responses were influenced by certain student demographic and socio-cultural factors. To ensure that the framing sample was representative of the target audience (Lang undergraduate and graduate students), students were recruited through various club associations and identified classes to include all cohorts and students with different extracurricular interests. Clubs and classes chosen for the study were identified as either sustainability-focused or not. The questionnaire was first distributed in the fall of 2020 and was then distributed a second time in winter 2021. In both cases, face-to-face classes and on-campus study were restricted due to COVID-19. A mobile-friendly application of the survey was added to leverage the ubiquity of smartphones to better access students in remote locations and to simplify the survey completion task. Recognizing the fatigue aspect associated with a continuous lockdown order, we coded questionnaires to capture this difference in timing (i.e., fall and winter). We highlight the potential impacts of COVID-19 on our survey in the discussion section.

The questionnaire consisted of 64 questions including twenty questions that form the PIR rating, assessing how students perceive their school’s current commitment to creating a positive impact, twenty-three socio-cultural and demographic questions, eleven attitude and behavior questions to establish their sustainability attitudes and behavior score, eight political questions to establish their political leaning score, and two overall satisfaction with their academic journey questions. Three treatments were conducted; the first treatment placed the PIR survey questions first and the second placed the PIR questions second to test whether priming or framing effects from the other questions would influence the score. Both treatments were conducted in the fall of 2020. The third treatment had the PIR questions second (same as treatment 2), but was executed in the winter of 2021. The electronic questionnaire took approximately 40 min to complete. A subset of the questions from the survey is attached in Appendix A.

There are three parts to this analysis. In Part 1, we estimate the reliability of the PIR scale using the coefficient alpha (also known as Cronbach's alpha) using R statistical programming language and the 'psy' package. Cronbach's alpha [48] measures the average covariance between the items in the scale, where the higher the coefficient alpha the more consistent the measure of the concepts that form the scale. Knekta et al. [49] recommended Cronbach's alpha to estimate reliability but not to estimate the validity for an instrument with a single distribution. Although a high reliability is necessary to make valid interpretations, it is not sufficient as a validity measure, as it does not test for the dimensionality of a scale. To test for the dimensionality of a scale, a factor analysis is recommended.

Therefore, in Part 2 of the analysis, a confirmatory factor analysis (CFA) was conducted to verify the factor structure of the 20-observed responses that informed the PIR scale. A CFA was chosen as it measures a model that is hypotheses-driven. Specifically, for the PIR, the researchers pre-specified all aspects of the model. Questions in the survey were divided into three areas, and these were further divided into seven dimensions. Yong and Pearce [50] and Kline [51] recommended the use of CFA for questionnaires that exceed 4 items and that have at least three or more items for each subcategory identified. In our case, the PIR has 20 items, and there are at least three items associated with each of the three subcategories of engaging, energizing, and educating. Although larger sample sizes are generally better to increase the statistical power of the result [52], minimum sample sizes of 100 to 200 can be utilized, as long as the free parameters are less than the known values, i.e., the model is over-identified [53].

The CFA tested the hypothesis that a relationship exists between the observed responses and their underlying latent constructs; specifically, the area of categories and dimensions. R statistical programming language and the 'lavaan' package were used to perform the CFA. A maximum likelihood estimation was chosen given normally distributed data. A covariance matrix explored the psychometric properties of the 20-item PIR survey. To determine model fit, we report the chi-square value, comparative fit index (CFI), the Tucker–Lewis fit index (TLI), and the root mean square error of approximation (RMSEA), where a $CFI \geq 0.90$, $TLI \geq 0.95$, and $RMSEA < 0.08$ would indicate a good fit.

In Part 3, we conducted a bivariate (OLS) statistical analysis to understand the causes of the observed PIR scores. A bivariate (OLS) versus multivariate statistical model was implemented even though there is a trend in multivariate analysis in this area [54,55], as the bivariate OLS model was a best fit given the theory and the research design for this study. Specifically, the model satisfied the 7 assumptions necessary for a linear regression to give a valid result; the dependent variable was continuous; the independent variables were categorical; there was a linear relationship between the dependent and independent variables; there were no significant outliers; there was independence of observations; the data were homoscedastic; and the residuals (errors) of the regression line were approximately normally distributed. The dependent and explanatory variables are described below.

4.1. Dependent Variables

The overall PIR score and scores of three subcategories of the PIR scale (energizing, educating, and engaging) were calculated following the original methodology in the PIR report. Specifically, the arithmetic average over 20 PIR questions for each participant was defined as the overall PIR score. Three sub-scores were calculated as the arithmetic average over related PIR questions.

Table 1 reports the summary statistics for the aforementioned four PIR scores. The average PIR score was 7.43, positioning the Lang Business School as a transforming (7.4–8.7) business school on the scale's tiered rating system. The scores for energizing and engaging subcategories were also transforming, with scores of 7.65 and 7.45, respectively. The educating subcategory scores were positioned lower at 7.29, placing Lang on the progressing (5.9–7.3) tier within the tiered system (see Appendix B, for details on PIR tiers)

Table 1. Summary statistics of PIR scores. This table reports the number of valid surveys and mean, standard deviation, minimum, maximum, and three quantile levels of the PIR score and its 3 subcategories.

	PIR	Energizing	Educating	Engaging
count	143	143	143	143
mean	7.43	7.65	7.29	7.45
std	1.32	1.33	1.44	1.48
min	4.30	3.83	3.30	3.00
25%	6.55	6.67	6.50	6.50
50%	7.50	7.83	7.40	7.50
75%	8.45	8.67	8.35	8.50
max	9.90	10.00	10.00	10.00

Treatment 3 (winter 2021) was not significantly different from either Treatment 1 or 2 (fall 2020). As such, Treatment 3 and 2 were collapsed, as they both had the PIR questions positioned second in the survey. Table 2 reports the two-sided *t*-test results on the equality of PIR and three subcategory scores from the two treatments. *p*-values greater than 10% suggested that there are no statistically significant differences between the two treatments.

Table 2. Two-sided *t*-test results on PIR and its 3 subcategories. This table reports the value of *t*-statistic, degree of freedom (DoF), *p* value, and 95% confidence interval.

Variable	<i>t</i> Statistic	DoF	<i>p</i> Value	CI 95%
PIR	0.87	141	0.39	[−0.25, 0.63]
Energizing	0.26	141	0.79	[−0.38, 0.50]
Educating	1.58	141	0.12	[−0.10, 0.85]
Engaging	−0.32	141	0.75	[−0.57, 0.41]

4.2. Explanatory Variables

Thirteen (13) explanatory variables were constructed from the 44-question survey to test the influences of survey design, survey distribution methods, and student demographic and, socio-cultural factors on PIR scores. A subset of variables were direct response categorical variables. These included a course that requested that you take the survey, whether you belonged to a student club or not and if so whether the club had a sustainability focus, self-identified faith affiliation, overall satisfaction with their academic experience, gender, subject discipline, and co-op status.

An additional three explanatory variables were constructed indirectly based on a series of questions. In the first step, a continuous index is constructed (see details below). In the second step, a binary variable is constructed based on whether the score is below or above the median score among all participants.

4.2.1. Political Leaning Index (Political)

This index was based on 8 questions from a pre-existing PEW survey, where a lower index value would suggest a more liberal-leaning (left-leaning) and a higher index value would suggest a more conservative-leaning (right-leaning) political view. As the political orientation of students was not significant with a low standard deviation, we were interested in understanding whether being relatively left-leaning versus relatively right-leaning influenced the score. To this end, we constructed a binary variable with two levels with the reference point as left-leaning (see questions 1–8, Appendix A)

4.2.2. Attitudes toward Sustainability and the Environment (Envir Belief)

An index score ranging from 1 to 5 was constructed per subject based on four questions related to their attitudes toward sustainability and the environment. The higher score indicated that the participant was more sustainability concerned. This index had similar

results as the political leaning index score, leading us to the construction of a binary variable with two levels with the reference point as ‘lower sustainability attitude’ (see questions 11, 15, and 16, Appendix A)

4.2.3. Consumption and Purchase Behavior (Consum)

A binary variable was created based on a series of four questions asking students about their consumption and purchase behavior. Zero was assigned to selected responses if the purchasing and intended consumption behavior did not comply with sustainability concerns (see questions 9 and 12–14, Appendix A).

5. Results

In total, we collected 156 PIR responses usable for reliability and validity testing and 143 usable surveys to conduct the bi-variant analysis. Table 3 reports the summary statistics for the political leaning score (political), attitudes toward sustainability and environment score (env_belief), and the consumption and purchase behavior score (consum). The mean score of 0.30 indicated that participants at Lang were politically left-leaning, aligning more closely to liberal government policies. The high mean of 4.17 out of a possible score of 5 indicated a positive attitude toward sustainable business practices and the environment. Conversely, the mean score of 0.42, below the median of 0.50, indicated an intended consumption and purchase behavior marginally away from environmentally sustainable products.

Table 3. Summary statistics of three indexes. This table reports the number of valid surveys and mean, standard deviation, minimum, maximum, and three quantile levels of political leaning index, environmental belief index, and shopping habit index.

	Political	Env Belief	Consum
count	143	143	143
mean	0.30	4.17	0.42
std	0.21	0.57	0.30
min	0.00	2.50	0.00
25%	0.12	3.75	0.33
50%	0.25	4.25	0.33
75%	0.38	4.50	0.67
max	0.88	5.00	1.00

Table 4 reports the descriptive statistics for PIR scores (mean and standard deviation). The majority of participants identified as female (65%), were under-graduates (77%), were in a non-co-op program (63%), did not belong to a club (75%), identified as no faith (45%), and had a ‘relatively’ left (liberal) political leaning (57%). Close to 49% of participants were recruited from a course with a sustainability focus. With the exception of the high number of females (38% versus 65%), the sample was representative of the target population at Lang [56].

Table 4. PIR scores by categories. This table reports the average PIR score and its standard deviation for each category. The number of valid surveys for each category is also reported. The first level within each category is the reference level in the regression.

Category	Mean	N	STD
Sustainability-Focused Course			
No	7.28	73	1.43
Yes	7.59	70	1.18
Gender			
Male	7.54	48	1.37
Choose not to respond	7.70	2	0.42
Female	7.37	93	1.31

Table 4. *Cont.*

Category	Mean	N	STD
Faith			
Faith	7.60	79	1.22
No Faith	7.22	64	1.42
Degree			
Undergraduate	7.52	110	1.27
Graduate	7.12	33	1.46
Co-Op			
No	7.24	90	1.37
Yes	7.75	53	1.18
Academic Evaluation			
Meet	7.03	46	1.20
Above	7.77	87	1.26
Below	6.28	10	1.30
Clubs Membership			
No	7.28	107	1.31
Yes	7.88	36	1.25
Sustainability-Focused Clubs Membership			
No	7.38	125	1.30
Yes	7.78	18	1.44
Political Leaning			
Left	7.52	82	1.31
Right	7.31	61	1.34
Environmental Belief			
Less	7.36	68	1.24
More	7.49	75	1.39
Consumption Habit			
Less	7.54	32	1.31
More	7.40	111	1.33
Treatment			
First	7.53	68	1.25
Second	7.34	75	1.38

All undergraduate participants were registered in one of 11 subprograms housed within the Bachelor of Commerce program. All graduate participants were registered as Lang Master's students. Table 5 reports the average PIR score for participants from different programs. The box plots in Figure 1 show the distribution of scores by academic program.

Table 5. Descriptive statistics by academic programs. This table reports the average PIR score and its standard deviation for different academic programs. The number of valid participants from each academic program is also reported. Specifically, Panel A contains all undergraduate programs: Bachelor of Commerce and others. Panel B contains all graduate programs.

Program	Mean	N	STD
Panel A: Undergraduate Programs Bachelor of Commerce			
Accounting	7.28	16	0.95
Food and Agricultural Business	6.18	3	0.62
Hotel Food and Tourism Management	7.29	15	1.40
Management	7.84	9	1.23
Management Economics and Finance	7.75	17	1.20
Marketing Management	7.70	24	1.41
Organizational Leadership	8.16	5	1.02
Public Management	7.65	5	0.36
Real Estate and Housing	9.18	4	0.53
Sport and Event Management	7.20	2	1.13
Undeclared	6.59	10	1.23

Table 5. Cont.

Program	Mean	N	STD
Panel B: Graduate Programs			
MA	7.05	4	1.74
MA (Leadership)	8.40	2	0.14
MBA (HFTM)	8.38	3	0.13
MBA (Sustainability)	7.20	12	1.55
MSc	6.52	12	1.36

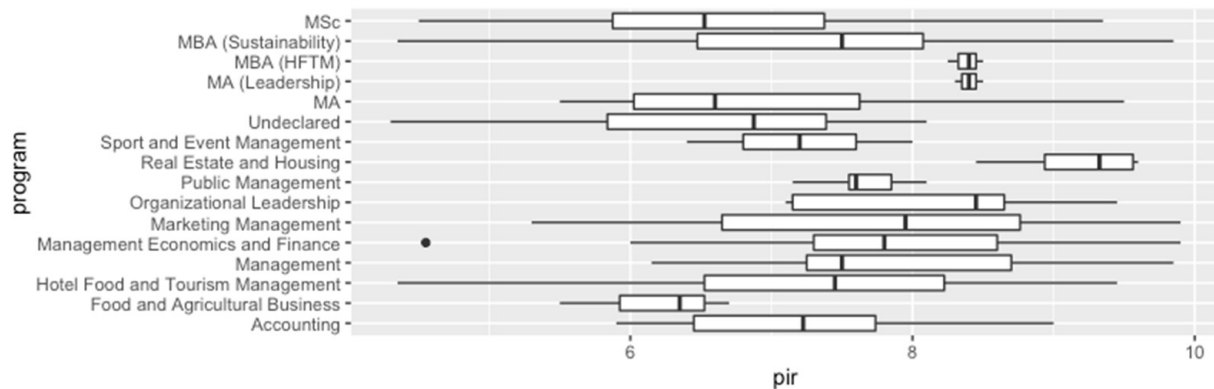


Figure 1. Box plots of PIR scores by academic discipline.

5.1. Coefficient Alpha Estimate

The Cronbach alpha calculated the average correlation between all 20 items contained in the survey. The coefficient alpha (α) was 0.949. In the absence of the same students repeating the survey multiple times, this high correlation provides a good measure of the scales' ability to provide consistent results (reliability). Thus, there is support for Hypothesis 1. The PIR is a reliable instrument as estimated by the coefficient alpha (Cronbach's alpha).

5.2. Confirmatory Factor Analysis

The CFA was first conducted using the latent variables (on the left in Tables 6 and 7) comprised of the indicators (observed variables on the right in Tables 6 and 7). These are the areas and dimensions with the associated questions as selected by the original creators of the PIR scale. The three model fit criteria for the CFA and coefficients can be found in Table 6. The chi-square value (p -value) = 0.00, the comparative fit index (CFI) = 0.835, the Tucker–Lewis fit index (TLI) = 0.812, and the RMSEA = 0.114 (see Table 8) The chi-square result rejected the null hypothesis that the model fits the data. The CFI, TLI, and RMSEA values also indicated a poor fit between the model constructs and the observed data.

Next, we conducted a CFA using the seven dimensions of governance, culture, programs, learning methods, student support, the institution as a role model, and public engagement to determine whether this led to a better fitting model. The results of the second analysis can be found in Tables 7 and 8. The chi-square (p -value) = 0.00, the comparative fit index (CFI) = 0.869, the Tucker–Lewis fit index (TLI) = 0.833, and the RMSEA = 0.107. These values again indicated a poor fit between the model constructs and the observed data.

Table 6. CFA with three factors. This table reports the estimated coefficients and standard errors using confirmatory factor analysis with three latent factors that are in the first column. The second column contains the 20 observable variables, which are reported using the corresponding question numbers in the PIR system.

Latent Variable	Observed Variable	Coefficients	SE
Energizing	Q1	1.13	0.11
	Q2	1.38	0.12
	Q3	1.33	0.12
	Q4	1.11	0.14
	Q5	1.43	0.13
	Q6	1.35	0.12
Educating	Q7	1.25	0.13
	Q8	1.26	0.15
	Q9	1.30	0.14
	Q10	1.40	0.15
	Q11	1.31	0.13
	Q12	1.44	0.14
	Q13	1.70	0.16
	Q14	1.46	0.14
	Q15	1.31	0.12
	Q16	1.48	0.14
Engaging	Q17	1.54	0.14
	Q18	1.44	0.13
	Q19	1.46	0.12
	Q20	1.27	0.13

Table 7. CFA with seven factors. This table reports the estimated coefficients and standard errors using confirmatory factor analysis with seven latent factors that are in the first column. The second column contains the 20 observable variables, which are reported using the corresponding question numbers in the PIR system.

Latent Variable	Observed Variable	Coefficients	SE
Governance	Q1	1.25	0.11
	Q2	1.56	0.12
Culture	Q3	1.36	0.12
	Q4	1.11	0.14
	Q5	1.45	0.13
	Q6	1.37	0.12
Program	Q7	1.32	0.13
	Q8	1.28	0.15
	Q9	1.30	0.14
	Q10	1.37	0.15
Learn	Q11	1.35	0.13
	Q12	1.47	0.14
	Q13	1.69	0.17
Support	Q14	1.66	0.13
	Q15	1.48	0.12
	Q16	1.41	0.15
Model	Q17	1.46	0.15
	Q18	1.34	0.14
Public	Q19	1.41	0.12
	Q20	1.25	0.13

Table 8. Confirmatory factor analysis result comparison. This table reports the three criteria in the confirmatory factor analysis: CFI, TLI, and RMSEA.

Model	CFI	TLI	RMSEA
Three Latent Factors	0.835	0.812	0.114
Four Latent Factors	0.869	0.833	0.107
Seven Latent Factors	0.862	0.831	0.119

These two analyses indicated that the original categorized survey questions may not be gathering the correct information to measure that pre-specified theme. Using the covariance matrix that explored the psychometric properties of the 20-item PIR scale (see Table 9), we constructed a new model by placing the responses with the highest covariances together to see if new latent variables emerged that could better explain the data. Specifically, we investigated whether the stronger covariance among items was potentially due to one common single factor. The covariance matrix informed a four-factor model (see Tables 8 and 10). The chi-square value (p -value) = 0.00, the comparative fit index (CFI) = 0.862, the Tucker–Lewis fit index (TLI) = 0.831, and the RMSEA = 0.119, which again indicated a poor fit. Thus, Hypothesis 2 is rejected. The confirmatory factor analysis does not support the PIR as a valid instrument.

Table 9. Correlation between 20 PIR questions.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19
Q2	0.71																		
Q3	0.54	0.66																	
Q4	0.33	0.38	0.53																
Q5	0.52	0.56	0.63	0.49															
Q6	0.53	0.57	0.62	0.47	0.57														
Q7	0.47	0.60	0.61	0.37	0.50	0.62													
Q8	0.23	0.34	0.43	0.38	0.39	0.45	0.42												
Q9	0.56	0.46	0.40	0.25	0.51	0.57	0.57	0.35											
Q10	0.35	0.42	0.48	0.39	0.52	0.42	0.44	0.45	0.44										
Q11	0.51	0.49	0.37	0.34	0.53	0.45	0.52	0.49	0.68	0.46									
Q12	0.34	0.43	0.36	0.42	0.42	0.51	0.45	0.65	0.46	0.45	0.55								
Q13	0.43	0.49	0.47	0.48	0.53	0.53	0.47	0.41	0.47	0.55	0.46	0.60							
Q14	0.44	0.48	0.39	0.40	0.53	0.56	0.36	0.47	0.45	0.44	0.49	0.62	0.61						
Q15	0.45	0.51	0.50	0.43	0.56	0.56	0.40	0.40	0.49	0.48	0.48	0.54	0.62	0.76					
Q16	0.45	0.56	0.57	0.45	0.52	0.57	0.55	0.44	0.46	0.48	0.50	0.49	0.43	0.55	0.57				
Q17	0.52	0.49	0.35	0.44	0.44	0.53	0.41	0.37	0.54	0.55	0.55	0.55	0.63	0.64	0.57	0.55			
Q18	0.47	0.53	0.52	0.45	0.49	0.60	0.51	0.34	0.37	0.54	0.39	0.42	0.45	0.45	0.44	0.58	0.52		
Q19	0.41	0.55	0.59	0.41	0.53	0.56	0.59	0.44	0.41	0.58	0.49	0.56	0.53	0.48	0.47	0.67	0.64	0.69	
Q20	0.52	0.57	0.46	0.40	0.44	0.48	0.51	0.30	0.50	0.39	0.47	0.41	0.41	0.48	0.45	0.45	0.55	0.60	0.56

Table 10. CFA with four factors. This table reports the estimated coefficients and standard errors using the confirmatory factor analysis with four latent factors that are in the first column. The second column contains the 16 observable variables, which are reported using the corresponding question numbers in the PIR system.

Latent Variable	Observed Variable	Coefficients	SE
Factor 1	Q1	1.26	0.11
	Q2	1.48	0.12
	Q20	1.24	0.13
Factor 2	Q6	1.31	0.12
	Q8	1.21	0.16
	Q9	1.41	0.14
	Q11	1.38	0.14

Table 10. Cont.

Latent Variable	Observed Variable	Coefficients	SE
Factor 3	Q12	1.43	0.14
	Q13	1.72	0.16
	Q14	1.51	0.13
	Q16	1.48	0.14
	Q17	1.58	0.14
	Q19	1.34	0.12
Factor 4	Q10	1.35	0.16
	Q15	1.30	0.12
	Q18	1.28	0.14

5.3. OLS Regression Analysis

OLS regressions with no interaction terms (Table 11) and with interaction terms (Table 12), with either the continuous PIR Score or the three PIR subdimension scores as the dependent variable, were explored to test if there is a selection bias in channeling the data collection for the PIR through student organizations engaged in the sustainability field, from students in courses linked to sustainability, and other demographic or socio-cultural characteristics.

Table 11. OLS regression without interaction terms. This table reports the results from four OLS regressions. The standard errors are reported in parenthesis. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

	Dependent Variable:			
	Pir	Pir_Energizing	Pir_Educating	Pir_Engaging
	−1	−2	−3	−4
Course (Y)	0.625 **	0.383	0.750 **	0.676 **
	−0.284	−0.284	−0.314	−0.331
Gender (No Response)	−0.045	−0.168	0.026	−0.037
	−0.899	−0.899	−0.992	−1.046
Gender (Female)	−0.110	−0.021	−0.109	−0.248
	−0.231	−0.231	−0.254	−0.268
Faith (N)	−0.426 *	−0.444 **	−0.393	−0.484 *
	−0.216	−0.216	−0.238	−0.251
Degree (Graduate)	−0.286	−0.209	−0.379	−0.168
	−0.342	−0.342	−0.378	−0.398
Co-op (Y)	0.318	0.469 *	0.283	0.177
	−0.241	−0.241	−0.266	−0.281
Club (Y)	0.508	0.486	0.41	0.789 **
	−0.333	−0.333	−0.367	−0.387
Sus Club (Y)	−0.131	−0.351	0.068	−0.299
	−0.415	−0.415	−0.458	−0.483
Eval (Above)	0.702 ***	0.826 ***	0.636 **	0.682 **
	−0.226	−0.226	−0.249	−0.262
Eval (Below)	−1.072 **	−0.759 *	−1.345 ***	−0.860 *
	−0.433	−0.433	−0.477	−0.503
Treatment (Second)	0.291	0.287	0.193	0.543
	−0.336	−0.336	−0.37	−0.39
Political (Right)	−0.011	−0.152	−0.018	0.219
	−0.23	−0.23	−0.253	−0.267
Belief (Sus)	0.409 *	0.519 **	0.363	0.358
	−0.231	−0.231	−0.255	−0.269
Shopping (Sus)	−0.494 *	−0.514 *	−0.523 *	−0.390
	−0.272	−0.272	−0.3	−0.317
Constant	6.892 ***	7.061 ***	6.870 ***	6.695 ***
	−0.424	−0.424	−0.468	−0.494

Table 11. Cont.

	Dependent Variable:			
	Pir	Pir_Energizing	Pir_Educating	Pir_Engaging
	−1	−2	−3	−4
Observations	143	143	143	143
R ²	0.259	0.267	0.245	0.199
Adjusted R ²	0.178	0.187	0.163	0.112
Residual Std. Error (df = 128)	1.196	1.196	1.319	1.391
F Statistic (df = 14; 128)	3.196 ***	3.328 ***	2.972 ***	2.277 ***

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 12. OLS regression with interaction terms. This table reports the results from four OLS regressions. The standard errors are reported in parenthesis. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

	Dependent Variable:			
	Pir	Pir_Energizing	Pir_Educating	Pir_Engaging
	−1	−2	−3	−4
Course (Y)	0.630 **	0.388	0.755 **	0.682 **
	−0.282	−0.282	−0.312	−0.328
Gender (No Response)	−0.204	−0.335	−0.117	−0.224
	−0.897	−0.896	−0.993	−1.043
Gender (Female)	−0.159	−0.072	−0.153	−0.305
	−0.23	−0.23	−0.255	−0.268
Faith (N)	−0.434 **	−0.452 **	−0.399 *	−0.493 **
	−0.214	−0.214	−0.237	−0.249
Degree (Graduate)	−0.277	−0.200	−0.371	−0.157
	−0.34	−0.339	−0.376	−0.395
Co-op (Y)	0.317	0.469 *	0.282	0.176
	−0.239	−0.239	−0.265	−0.278
Club (Y)	0.514	0.492	0.415	0.796 **
	−0.33	−0.33	−0.366	−0.384
Sus Club (Y)	−0.123	−0.343	0.075	−0.290
	−0.412	−0.411	−0.456	−0.479
Eval (Above)	0.707 ***	0.831 ***	0.641 **	0.688 ***
	−0.224	−0.223	−0.248	−0.26
Eval (Below)	−1.110 **	−0.798 *	−1.379 ***	−0.904 *
	−0.43	−0.429	−0.476	−0.5
Treatment (Second)	0.264	0.259	0.168	0.511
	−0.333	−0.333	−0.369	−0.388
Political (Right)	0.352	0.229	0.308	0.646 *
	−0.308	−0.308	−0.342	−0.359
Belief (Sus)	0.743 **	0.869 ***	0.664 **	0.751 **
	−0.299	−0.298	−0.331	−0.347
Shopping (Sus)	−0.476 *	−0.496 *	−0.507 *	−0.370
	−0.27	−0.27	−0.299	−0.314
Political (Right) * Belief (Sus)	−0.777 *	−0.815 *	−0.699	−0.915 *
	−0.446	−0.445	−0.493	−0.518
Constant	6.691 ***	6.850 ***	6.688 ***	6.458 ***
	−0.437	−0.436	−0.483	−0.508
Observations	143	143	143	143
R ²	0.276	0.286	0.257	0.219
Adjusted R ²	0.191	0.201	0.169	0.126
Residual Std. Error (df = 127)	1.187	1.185	1.314	1.38
F Statistic (df = 15; 127)	3.234 ***	3.387 ***	2.930 ***	2.368 ***

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

The benchmark model is Model (1) from Table 11. The results of a multiple linear regression showed a collective significant effect of all the independent variables, $F(14, 128) = 3.196$, and $R^2 = 0.259$. Specifically, 25.9% of the variance was explained by the model. Sustainability-focused courses ($\beta = 0.625, t = 2.199, p = 0.030$), academic evaluation above expectations ($\beta = 0.702, t = 3.113, p = 0.003$), and attitudes toward the environment ($\beta = 0.409, t = 1.770, p = 0.080$) were positive and significant in the model, while identifying with no faith ($\beta = -0.426, t = -1.978, p = 0.051$), academic evaluation below expectations ($\beta = -1.072, t = -2.477, p = 0.015$), and consumption behavior ($\beta = -0.494, t = -1.813, p = 0.073$) were negative and significant in the model. Students who were requested to complete the survey within a course that taught sustainability topics, students who rated their academic experience as exceeding expectations, and students who had a positive attitude toward the environment had higher PIR scores. Conversely, students who identified with 'no faith', students who had an academic evaluation below expectations, and students who identified with lower eco-conscious consumption and purchase behavior had lower PIR scores.

To study the effect of the explanatory variables on the three subdimensions of the PIR system, three more OLS regressions were run with the three PIR subdimension scores as the dependent variables. The general effect of the explanatory variables on these were similar to that on the general PIR, with several significant differences. Firstly, whether participants were from sustainability-focused courses had no significant effect on the energizing dimension ($\beta = 0.383, t = 1.348, p = 0.181$). On the other hand, the energizing dimension was the only one that was significantly affected by whether the participants were from co-op programs ($\beta = 0.469, t = 1.946, p = 0.054$). Secondly, faith and sustainability attitudes had no significant effect on educating. Thirdly, club membership had a significant and positive influence on the engaging score ($\beta = 0.789, t = 2.038, p = 0.044$).

In Table 12, the interaction term between political leaning and environmental belief was negative and significant, except for Model (3). On average, participants with a political vision leaning to the right (aligned with conservative policies) and a more sustainability-focused environmental belief would have a significantly lower PIR score ($\beta = -0.777, t = -1.744, p = 0.084$). The magnitude of the influence from this interactive term on the energizing dimension ($\beta = -0.815, t = -1.832, p = 0.070$) and the engaging dimension ($\beta = 0.915, t = -1.767, p = 0.080$) was similar. The effect on the educating dimension was negative but not statistically significant ($\beta = 0.699, t = -1.416, p = 0.160$).

Hypothesis 3 was rejected; there was no selection bias in channeling the data collection for the PIR through student organizations engaged in the sustainability field. There was support for Hypotheses 4 and 5. There was a selection bias in collecting PIR data from students in courses linked to sustainability (H4). Demographics and socio-cultural characteristics of the student influenced PIR responses (H5).

6. Discussion

The PIR scale is a promising scale for selection by schools like the Lang School of Business who 'are committed to using business as a force for good to achieve the United Nation's SDGs' [56]. In addition to providing a benchmark for a business school's performance as perceived by students, arguably its most important stakeholder, it is a tool that could help 'attract students and faculty who have a social conscience, an environmental sensibility, and a commitment to community involvement' [56]. Building an organizational culture that is aligned with the mission, vision, and core values of the institution is critical to achieving an organization's intended goals. Given the perverse effects that traditional published ranking scales can cause, careful consideration is needed to ensure alignment. Confirming the reliability and validity of any chosen scale is essential. The PIR provides transparency in both the criteria used and methodologies employed. The creators are committed to developing a scale that helps the business community (including the academic community) realize the role it plays in ensuring a sustainable future for all stakeholders.

Given the power of published ratings and the intention of the PIR scale, we identify areas for consideration and improvement toward a statistically robust PIR scale and an execution strategy for the survey that could help mitigate unintended biases.

Firstly, the survey's coefficient alpha is high, indicating the scale's reliability, yet the CFA analysis pointed to an invalid model. A survey's coefficient alpha can be high and yet the survey instrument could still not be measuring what the researcher intended to measure [57–60]. In this case, the CFA analysis revealed that all questions that informed the survey were highly interrelated. Specifically, the observed responses for the latent variables (i.e., energizing, educating, and engaging as well as the seven dimensions) were too interconnected and were not separate enough to clearly measure three distinct themes or seven separate dimensions. Tables 13 and 14 highlight the high correlations between the variables.

Table 13. Correlation between three latent factors. this table reports the correlation between three subcategories in the PIR system.

	Energizing	Educating
Educating	0.79	
Engaging	0.74	0.79

Table 14. Correlation between seven latent factors. This table reports the correlation between seven subcategories in the PIR system.

	Governance	Culture	Program	Learn	Support	Model
Culture	0.68					
Program	0.60	0.73				
Learn	0.58	0.67	0.79			
Support	0.60	0.71	0.68	0.74		
Model	0.62	0.67	0.68	0.69	0.72	
Public	0.63	0.67	0.69	0.65	0.66	0.8

This inter-relatability between all questions suggested that a better fit model could be a one-factor model with 20 indicators. However, the CFA results for the one-factor model indicated a poor fit (chi-square = 0.00, CFI = 0.81, TLI = 0.78, RMSEA = 0.12), suggesting room for improvement.

However, the results of these CFA analyses are contestable on two fronts. First, the seven-factor CFA analysis did not meet the specifications for CFA analysis. Specifically, for some of the seven categories, there only two questions, and it is recommended that more than two questions are needed for each identified theme for patterns to occur. Second, although a CFA analysis is applicable to small samples (156) where the free parameters are less than the known values (over-identified), CFA and general class structural equation models are large sample techniques. Therefore, the larger the sample the better. Kline [51] recommends the N:q rule, specifically, that the sample size should be determined by the number of q parameters in the model and that rule should be 20:1. In our example, this would suggest a more valid sample size of approximately 1200.

To this end, our first recommendation is to conduct a CFA analysis with a larger data set to corroborate these initial findings. If the subsequent CFA shows similar results to the ones found in this study, we recommend a revision of the survey questions ensuring that the questions associated with each identified theme have a high covariance within each category and a lower covariance between the selected categories, indicating the measurement of distinct themes or concepts. Distinct themes help inform the participating institution on explicit areas to focus on for improvement. Additionally, if the CFI, TLI, and RMSEA results from the larger data set fail to reject the null, indicating it is not a bad model, we still cannot necessarily say it is the best model. Therefore, using the larger data set, we would further recommend testing for other latent constructs that may have emerged.

Further, we recommend an exploratory factor analysis (EFA) or a clustering algorithm from an unsupervised machine learning technique to explore patterns underlying the data set to help develop new theories and identify items that do not empirically belong and should consequently be removed from the survey.

Survey distribution methods and socio-cultural factors influenced student PIR scores. Survey distribution was not completely randomized. A subset of faculty was selected who would be willing to request students to complete the survey, and leaders of extra-curricular school-sanctioned clubs were asked to distribute the survey to their members. Students who were requested to take the survey through a class that taught sustainability and/or corporate responsibility topics had significantly higher PIR scores versus students who were asked by a professor of a course that did not teach these topics.

Students who evaluated their academic experience at Lang as 'exceeding expectations' had a higher PIR score than students who rated their experience as 'meets' or 'below' expectations. Although belonging to a student club was not significant, students who belonged to a club were more likely to select that their academic experience 'exceeds expectations.' Interestingly, this observation suggests that if the published rating attracts students who are aligned with the goals of the institution, and the institution does not live up to the student expectations, then the subsequent scores will be lower. The best way forward, therefore, is to have a high rating that is a true representation of the student experience, as this will lead to subsequent high ratings. These initial findings suggest that the student-driven survey, properly disseminated, has a built-in mechanism toward continuous improvement.

The significant influence of the survey response 'academic experience exceeded expectation' on the PIR score and the correlation of this factor with students belonging to clubs require further unpacking. The dominant theoretical framework in general education literature suggests that extracurricular activity (ECA) (i.e., belonging to a club) has a positive impact on academic performance [61]. This literature indirectly connects higher academic performance with higher PIR scores. Establishing a direct and causal relationship between these two variables, in particular, that a higher PIR score signals higher academic performance by the students, could provide further benefits for schools who wish to participate in the rating and for employment recruiters of graduate students.

This study also tested explicitly for priming effects. In one survey treatment, socio-cultural, attitudinal, and political views were asked first before the PIR survey questions, and in the second treatment, these questions were asked in reverse. Although there was no significant difference in PIR scores between the two treatments, we cannot rule out a priming effect for students who were asked by a course instructor who teaches sustainability topics. Considerable experiments have shown how priming effects influence the behaviors of individuals [62–65].

Questions were included in the survey to identify faith affiliation, sustainable purchase and consumption behavior, and political orientation. These questions were included to understand the influence of pre-established North American values on North American business school PIR scores. It is important for subsequent studies that wish to test the influence of pre-established values that the questions change to reflect the situational context of the different geographic/political social environments in which the study is executed. At Lang (Guelph, ON, Canada), students were mainly left-leaning (liberal), and political orientation had no impact on PIR scores. However, those that identified with 'no faith' had a lower PIR than students who identified with faith. Students with higher environmental beliefs in terms of consumption and purchase behavior also had a higher PIR score. Literature has shown that sociocultural attributes could lead to biased results of surveys [66–69]. Although sociocultural differences are assumed in research involving humans, the results can be interpreted wrongly if there is no comparability [66].

One idea for consideration given these results is to include a set of pre-established value questions (non-political, non-religion-based) in the PIR that assess the organizational culture (OC) of the student body. For example, 'students in my program are more collabo-

rative than competitive.' Not only does this allow the institutions to test the alignment of the OC with its core values, but it also allows students to identify a school more closely aligned with their own values. This criteria for selection could continuously build student bench strength that allows a business school to deliver against its aligned vision.

6.1. Contributions to Knowledge and Literature

This study is phase one of a broader study toward uncovering embedded biases and incongruencies in methodological data collection procedures within business school ratings and rankings. While past research has been focused on uncovering embedded biases and methodological flaws in what many have described as 'rankings that cause perverse effects', for this first study, we chose a proactive approach by focusing on a new scale that is representative of student school experiences and a school's ability to have a positive impact on society beyond contributing to an organization's profits and GDP of a nation.

The scale was developed by academics, business school administrators, and students through a 'multi-step proto-typing process' [3] (p. 6). The PIR was first executed in 2019 to 2500 students, and the results of the survey were published in 2020 [3]. In its first execution, student organizations such as Oikos International, Net Impact, AIESEC, the SOS (Students Organising for Sustainability, in the UK), and Studenten voor Morgen (in The Netherlands) solicited students to complete the questionnaire [3]. Dyllick and Muff 'raised the question of selection bias' given that the student organizations executing the survey "are mainly active in the sustainability area" [3] (p. 9). They further highlight that "one of the participating schools has decided to do a controlled study of this question, based on their PIR 2020 data" [3] (p. 9). As the identified business school, we commend the PIR organization for inviting such a rigorous arms-length assessment and would encourage other ranking scales to follow suit. While we did not find selection biases from the execution of the study through student organizations, we did find that students participating in a sustainability-focused course, when asked by their instructor, had statistically significant higher PIR scores. The power influence of instructors may have played a role here; however, this would require further analysis. The reliability and validity tests, although not conclusive, the selection bias results, and the sociocultural influences on PIR score serve to provide direction toward enhancements to a rating scale that has the power to change the role that business schools play within society.

6.2. Implications for Business School Leaders and Business Practitioners

Business school ratings and rankings serve a dual purpose. Firstly, business school rankings signal to the community whether the metrics are legitimate or not, and how the school is performing in comparison to others and therefore serve as a powerful student recruitment tool. Secondly, it has the potential to influence the strategic direction of a school and the priorities of the faculty. The former is driven by an external audience and is influenced wrongly or rightly through media and a general acceptance by business schools as a crowning achievement. The second one implies careful consideration of the right 'measurement tool' that ensures performance of the organization that moves them toward the intended goals. These two purposes should be aligned and arguably in reverse order. Selecting the correct ranking and rating system to benchmark the organizational performance and ensuring a more valid and accurate ranking system serve to enhance institutional legitimacy by promoting behaviors internally that align with the school's vision, core values, and strategy.

For business practitioners who prefer to recruit from top-ranked business schools, understanding the criteria that underpin the traditional ranking scales is essential. In many cases, these criteria have no connection to the caliber of the student learning experience or the vision or core values espoused by the institution.

6.3. Limitations and Future Research Suggestions

The research was conducted during a global pandemic; it is difficult to determine how a student's assessment of the school's positive impact would be influenced by shut-down orders that forced them into an alternative learning environment. While upper-year students would have an institutional memory of lived experience on campus, first-year students would not. However, when comparing the PIR scores from this study with Lang's first-time participation scores (2019), we found no significant differences in the rating. With the exception of the number of females completing the survey (65% versus 38%), the sample was representative of the Lang student body. Although the sample was larger than the number of students that participated in the first PIR report (2020), it only represented 4% of the Lang student population.

Future research suggestions for the PIR scale specifically include: (1) conducting a confirmatory factor analysis (CFA) on a larger data set to determine the latent structure equation of the survey; (2) even if the subsequent CFA rejects the null, we recommend conducting an EFA to explore potential themes that may have been missed and to drop questions that are not empirically supported; (3) identifying an additional set of potential questions for consideration that measure student values; and (4) an additional same study at another business school in close proximity to Lang with traditional business school values to observe PIR differences to enhance the validity of the scale's ability to measure a school's positive social impact. This study is phase one of a broader research study that looks at methodological incongruencies and biases of 'published rankings', in particular rankings that influence the priorities of business schools, such as the FT50.

6.4. Conclusions

Published rankings, although beneficial for student recruitment, have caused unintended consequences. Perverse effects on faculty and decanal priorities have been identified, including incentivizing behaviors that are at odds with the achievement of the United Nations Sustainable Development Goals (SDGs) [1,2] and 2030 agenda. The Positive Impact Rating scale was introduced in response to these observations with the stated purpose to "help speed up the transformation" towards orienting "teaching, research and outreach activities towards social impact and sustainability" [3] (p. 6). For the PIR to achieve its goals, including becoming broadly perceived as a reliable assessment instrument, it is essential that its metrics and approach be held to a high standard and scale be statistically supported. The reliability of the scale was confirmed by the coefficient alpha. The validity of the scale, although not conclusive, requires further analysis with a larger data set with a recommendation to apply other structural equation modeling analysis. There were selection biases when distributing the scale and socio-cultural factors that influenced PIR scores, and these should be recognized when disseminating the survey and analyzing its results. Although these enhancements and considerations would improve the efficacy of the scale, the PIR scale continues to be a promising entrant to the 'published ranking' forum.

Author Contributions: Conceptualization, K.R. and J.C.H.; methodology, K.R.; validation, K.R., J.C.H. and T.R.; formal analysis, R.L. and K.R.; investigation, K.R. and T.R.; data curation, K.R., R.L. and T.R.; writing—original draft preparation, all four authors; writing—review and editing, K.R. and J.C.H.; visualization, R.L.; supervision, K.R.; project administration, K.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The members of the University of Guelph, Ontario, Canada Research Ethics Board have examined the protocol which describes the participation of the human participants in the above-named research project and considers the procedures, as described by the applicant, to conform to the University's ethical standards and the Tri-Council Policy Statement, 2nd Edition. REB number 20-06-033.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data available on request due to Research Ethic Board privacy restrictions given the nature of certain questions contained within the questionnaire.

Conflicts of Interest: One of the authors, Julia Christensen Hughes, is a former dean, who prioritized becoming “ranked” as part of her efforts to build the global brand of her business school. Following an analysis of various rankings, she identified Corporate Knights Better MBA ranking as the one most aligned with the aspirations of her school to “develop leaders for a sustainable world”. She also contributed to the development of the Positive Impact Rating and briefly served as a member of its board. At the invitation of Corporate Knights and the United Nations (UN) Global Compact and Principles for Responsible Management Education (PRME) initiative, Julia also facilitated several “Deans Dialogue” events at Davos during the World Economic Forum on business school rankings. More recently, she has engaged with the UN’s Higher Education Sustainability Initiative (HESI), through which she has continued to advocate for change in traditional rankings. While she provided input to the design of the current survey and its implementation, as well as discussions on the analysis of the results, to guard against any potential bias, she had no direct contact with the data or student participants.

Appendix A

Survey Questions Informing Explanatory Variables

Political leaning (Political), environmental beliefs (Environ Belief). And consumption and purchase behavior (Consum) help us assess students’ political leaning, attitudes toward the environment and consumption, and purchase behavior.

Please select the statement with which you agree the most [70].

1. A. Government is neither very effective nor efficient in getting the job done
B. Government often does a better job than people give them credit for
2. A. Government regulation of business is necessary to protect the public interest
B. Government regulation of business usually does more harm than good
3. A. People living below the poverty line in Canada today have it easy because they can get government benefits without doing anything in return
B. People living below the poverty line in Canada have hard lives because government benefits don’t go far enough to help them live decently
4. A. The Government should do more to help needy Canadians even if it means going deeper into debt
B. The Government today can’t do much more to help the needy
5. A. The best way to ensure peace is through military strength
B. Good diplomacy is the best way to ensure peace
6. A. Most corporations make a fair and reasonable amount of profit
B. Most corporations make too much profit
7. A. The amount of tax paid by corporations in Canada is fair and reasonable
B. Canadian Corporations should pay more tax
8. A. Stricter environmental laws and regulations cost too many jobs and hurt the economy
B. Stricter environmental laws are worth the cost
9. Please select the option that best describes your dietary choices.
 - a. Vegan (I Only eat plants and plant-sourced products)
 - b. Vegetarian (I abstain from the consumption of meat, shellfish, fish etc., but consume dairy, eggs)
 - c. Pescatarian (I am a vegetarian who does not eat meat but does eat fish and other shellfish)
 - d. Flexitarian (My diet is centered around plant foods, with the occasional inclusion of meat)
 - e. Omnivore (I eat all types of food including plants, meat, fish, shellfish, etc.)
 - f. Other

10. If you answered A, B, C, D, or F above, what is the reason for your food choice? Select the primary choice or all that apply.
 - a. Financial reasons
 - b. Health reasons
 - c. Ethical treatment of animal reasons
 - d. Environmental reasons
 - e. Religious reasons
 - f. other
11. Which of the following statements best describes what you believe to be true?
 - a. The climate is changing and human activity is mainly responsible
 - b. The climate is changing and human activity is partly responsible along with other factors
 - c. The climate is changing but human activity is not responsible at all
 - d. The climate is not changing
 - e. Don't know
12. Which of the following statements is true about you? Please select all that apply to you.
 - a. I drive an electric car
 - b. I plan to purchase an electric car in the near future
 - c. I closely monitor and try to minimize the energy I consume at home
 - d. I closely monitor and try to minimize the water I consume at home
 - e. I am signed up to an energy provider that focuses on renewable power sources
 - f. None of these
13. Which of the following statements are true about you when purchasing groceries? Please select all that apply to you.
 - a. prioritize products that are "locally sourced"
 - b. I prioritize products with a "fair trade" label
 - c. I prioritize products that are organic or GMO free
 - d. I prioritize products based on nutritional value
 - e. None of these
14. When deciding to buy a new (non-grocery) product or service, which attribute of the product or service most strongly influences your choice?
 - a. Price of the product or service
 - b. Availability of the product or service
 - c. The organization's (which produces the product or service) demonstrated commitment to the environment
 - d. The organization's (which produces the product or service) demonstrated commitment to social causes
 - e. The credibility of the brand in terms of product or service quality
 - f. Other

To what extent do you agree or disagree with the following statements. 7-point scale: strongly agree–strongly disagree.

15. The only purpose of a business is to provide a return on investment to its shareholders (e.g., maximize profits for the benefit of owners).
16. Business has a responsibility to all stakeholders, balancing shareholder return with the needs of employees, suppliers, customers, society at large, and the environment.
17. Increasing environmental disasters, for example, hurricanes, bush fires, and extreme hot and cold weather patterns, are due to climate change.

Appendix B

Table A1. PIR Tiers.

Level	Range	Difference	Characterization
Level 1	1–4.2		Beginning
Level 2	4.3–5.8	1.5 points	Emerging
Level 3	5.9–7.3	1.4 points	Progressing
Level 4	7.4–8.7	1.3 points	Transforming
Level 5	8.8–10	1.2 points	Pioneering

Source: 2020 Positive Impact Rating Report <https://www.positiveimpactrating.org/pir2020> (assessed on 8 March 2021).

References

- Badat, S. Global rankings of universities: A perverse and present burden. In *Global Inequalities and Higher Education: Whose Interests Are We Serving*; Entwistle, N., King, R., Eds.; Palgrave Macmillian: London, UK, 2010; pp. 117–141.
- Pitt-Watson, D.; Quigley, E. Business School Rankings for the 21st Century. UN Principles for Responsible Management Education (PRME). January 2019. Available online: <https://gmac.gcs-web.com/static-files/70eb162a-685b-4a83-81ae-748543aa3ad5> (accessed on 30 March 2021).
- Dyllick, T.; Muff, K. A Positive Impact Rating for Business Schools: Case Study. *Sustainability* **2020**, *12*, 9551. [CrossRef]
- Mio, C.; Panfilo, S.; Blundo, B. Sustainable development goals and the strategic role of business: A systematic literature review. *Bus. Strat. Environ.* **2020**, *29*, 3220–3245. [CrossRef]
- Dialogue at Davos: On Business School Rankings (23 January 2019) at the Invitation of Corporate Knights and Principles for Responsible Management Education. Available online: <https://grii.org/wp-content/uploads/2019/02/Davos-Dialogue-11-Feb-2019-FINAL.pdf> (accessed on 28 April 2021).
- Dyllick, T. Responsible management education for a sustainable world. *J. Manag. Dev.* **2015**, *34*, 16–33. [CrossRef]
- Christensen Hughes, J. Deans Multi-Stakeholder Dialogue at Davos: The Future of Business School Rankings and Ratings. PRME/Corporate Knights/Positive Impact Rating for Business Schools. 2020. Available online: <https://issuu.com/oikos-world/docs/davosrecapreport2020> (accessed on 9 April 2021).
- United Nations. Transforming Our World: The 2030 Agenda for Sustainable Development. 2015. Available online: <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication> (accessed on 28 April 2021).
- Khurana, K. *From Higher Aims to Hired Hands*; Princeton University Press: Princeton, NJ, USA, 2007.
- Peters, K. Business school rankings: Content and context. *J. Manag. Dev.* **2007**, *26*, 49–53. [CrossRef]
- Tyson, G.; Greif, G. A Look Behind the Numbers. In *Selections*; Graduate Management Admission Council: Reston, VA, USA, 2001; Volume 1, pp. 8–15.
- Elbeck, M. Advancing the relationship between business school ranking and student learning. *J. High. Educ. Policy Manag.* **2009**, *31*, 81–88. [CrossRef]
- Wilkins, S.; Huisman, J. UK business school rankings over the last 30 years (1980–2010): Trends and explanations. *High. Educ.* **2012**, *63*, 367–382. [CrossRef]
- Moon, H. The Globalization of Professional Management Education, 1881–2000: Its Rise, Expansion, and Implications. Ph.D. Thesis, Stanford University, Stanford, CA, USA, 2002. Unpublished.
- Bastedo, M.N.; Bowman, N.A. U.S. News & World Report College Rankings: Modeling Institutional Effects on Organizational Reputation. *Am. J. Educ.* **2010**, *116*, 163–183. [CrossRef]
- Elsbach, K.D.; Kramer, R.M. Members' Responses to Organizational Identity Threats: Encountering and Countering the Business Week Rankings. *Adm. Sci. Q.* **1996**, *41*, 442–476. [CrossRef]
- Espeland, W.N.; Sauder, M. Rankings and Reactivity: How Public Measures Recreate Social Worlds. *Am. J. Sociol.* **2007**, *113*, 1–40. [CrossRef]
- Sauder, M.; Espeland, W.N. The Discipline of Rankings: Tight Coupling and Organizational Change. *Am. Sociol. Rev.* **2009**, *74*, 63–82. [CrossRef]
- Sauder, M.; Fine, G.A. Arbiters, Entrepreneurs, and the Shaping of Business School Reputations. *Sociol. Forum* **2008**, *23*, 699–723. [CrossRef]
- Athavale, M.; Bott, J.; Myring, M.; Richardson, L. Deans' perceptions of published rankings of business programs. *J. Educ. Bus.* **2017**, *92*, 81–88. [CrossRef]
- Hallonsten, E. The impact of league tables and ranking systems on higher education decision making. *High. Educ. Manag. Policy* **2007**, *19*, 1–24.
- Hazelkorn, E. Learning to Live with League Tables and Ranking: The Experience of Institutional Leaders. *High. Educ. Policy* **2008**, *21*, 193–215. [CrossRef]
- Dasdan, A.; Van Lare, E.; Zivaljevic, B. How Reliable are University Rankings? *arXiv* **2020**, arXiv:2004.09006.

24. Mingers, J.; Willmott, H. Taylorizing business school research: On the ‘one best way’ performative effects of journal ranking lists. *Hum. Relat.* **2013**, *66*, 1051–1073. [[CrossRef](#)]
25. Rynes, S.L.; Giluk, T.L.; Brown, K. The Very Separate Worlds of Academic and Practitioner Periodicals in Human Resource Management: Implications for Evidence-Based Management. *Acad. Manag. J.* **2007**, *50*, 987–1008. [[CrossRef](#)]
26. Adler, N.; Harzing, A.-W. When Knowledge Wins: Transcending the Sense and Nonsense of Academic Rankings. *Acad. Manag. Learn. Educ.* **2009**, *8*, 72–95. [[CrossRef](#)]
27. Lee, N. Reflections on assessing academic quality in marketing, and the UK REF. *Eur. J. Mark.* **2011**, *45*, 477–483. [[CrossRef](#)]
28. Macdonald, S.; Kam, J. Ring a Ring o? Roses: Quality Journals and Gamesmanship in Management Studies. *J. Manag. Stud.* **2007**, *44*, 640–655. [[CrossRef](#)]
29. Macdonald, S.; Kam, J. Quality journals and gamesmanship in management studies. *Manag. Res. News* **2008**, *31*, 595–606. [[CrossRef](#)]
30. Bennis, W.G.; O’Toole, J. How business schools lost their way. *Harv. Bus. Rev.* **2005**, *83*, 96–104. [[PubMed](#)]
31. Dearden, J.A.; Grewal, R.; Lilien, G.L. Strategic Manipulation of University Rankings, the Prestige Effect, and Student University Choice. *J. Mark. Res.* **2019**, *56*, 691–707. [[CrossRef](#)]
32. Hall, J.; Martin, B.R. Towards a taxonomy of research misconduct: The case of business school research. *Res. Policy* **2019**, *48*, 414–427. [[CrossRef](#)]
33. Johnson, S.; Orr, K. What is business school research for? Academic and stakeholder perspectives, politics and relationality. *Stud. High. Educ.* **2020**, *45*, 557–578. [[CrossRef](#)]
34. Skapinker, M. Why Business Ignores the Business Schools. Financial Times. 8 January 2008. Available online: <https://www.ft.com/content/2198d908-280f-11e0-8abc-00144feab49a> (accessed on 15 March 2021).
35. Pfeffer, J.; Fong, C.T. The End of Business Schools? Less Success Than Meets the Eye. *Acad. Manag. Learn. Educ.* **2002**, *1*, 78–95. [[CrossRef](#)]
36. The Independent. The Complete University Guide: University League Table 2010. Available online: <http://www.independent.co.uk/news/education/higher/the-complete-university-guide-universityleague-table-2010-xhtml-1675616.html> (accessed on 15 March 2021).
37. Wedlin, L. *Ranking Business Schools: Forming Fields, Identities and Boundaries in International Management Education*; Edward Elgar Publishing: Gloucestershire, UK, 2006.
38. Cutter, C.; Elite, M.B.A. Programs Report Steep Drop in Applications. *Wall Street Journal*. 15 October 2019. Available online: <https://www.wsj.com/articles/elite-m-b-a-programs-report-steep-drop-in-applications-11571130001> (accessed on 5 February 2021).
39. Schlegelmilch, B.B. Why Business Schools Need Radical Innovations: Drivers and Development Trajectories. *J. Mark. Educ.* **2020**, *42*, 93–107. [[CrossRef](#)]
40. Morgeson, F.P.; Nahrgang, J.D. Same as It Ever Was: Recognizing Stability in the BusinessWeek Rankings. *Acad. Manag. Learn. Educ.* **2008**, *7*, 26–41. [[CrossRef](#)]
41. Cornuel, E. The role of business schools in society. *J. Manag. Dev.* **2005**, *24*, 819–829. [[CrossRef](#)]
42. Weybrecht, G. *The Sustainable MBA: A Business Guide to Sustainability*, 2nd ed.; John Wiley & Sons: West Sussex, UK, 2013.
43. Edwards, M.; Brown, P.; Benn, S.; Bajada, C.; Perey, R.; Cotton, D.; Jarvis, W.; Menzies, G.; McGregor, I.; Waite, K. Developing sustainability learning in business school curricula—Productive boundary objects and participatory processes. *Environ. Educ. Res.* **2020**, *26*, 253–274. [[CrossRef](#)]
44. Storey, M.; Killian, S.; O’Regan, P. Responsible management education: Mapping the field in the context of the SDGs. *Int. J. Manag. Educ.* **2017**, *15*, 93–103. [[CrossRef](#)]
45. Sroufe, R. Business Schools as Living Labs: Advancing Sustainability in Management Education. *J. Manag. Educ.* **2020**, *44*, 726–765. [[CrossRef](#)]
46. Hay, R.; Eagle, L. Impact of integrated sustainability content into undergraduate business education. *Int. J. Sustain. High. Educ.* **2020**, *21*, 131–143. [[CrossRef](#)]
47. Corporate Knights. Available online: <https://www.corporateknights.com> (accessed on 4 April 2021).
48. Cronbach, L.J. Coefficient alpha and the internal structure of tests. *Psychometrika* **1951**, *16*, 297–334. [[CrossRef](#)]
49. Knekta, E.; Runyon, C.; Eddy, S. One Size Doesn’t Fit All: Using Factor Analysis to Gather Validity Evidence When Using Surveys in Your Research. *CBE—Life Sci. Educ.* **2019**, *18*, rm1. [[CrossRef](#)] [[PubMed](#)]
50. Yong, A.G.; Pearce, S. A Beginner’s Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutor. Quant. Methods Psychol.* **2013**, *9*, 79–94. [[CrossRef](#)]
51. Kline, R.B. *Principles and Practice of Structural Equation Modeling*; Guilford Press: New York, NY, USA, 2015; ISBN 978-1-60623-876-9.
52. Gagne, P.; Hancock, G.R. Measurement Model Quality, Sample Size, and Solution Propriety in Confirmatory Factor Models. *Multivar. Behav. Res.* **2006**, *41*, 65–83. [[CrossRef](#)]
53. Whittaker, T.A.; Worthington, R.L. Item Response Theory in Scale Development Research. *Couns. Psychol.* **2016**, *44*, 216–225. [[CrossRef](#)]
54. O’Fallon, M.J.; Butterfield, K.D. Moral Differentiation: Exploring Boundaries of the “Monkey See, Monkey Do” Perspective. *J. Bus. Ethic.* **2011**, *102*, 379–399. [[CrossRef](#)]
55. Randall, D.M.; Gibson, A.M. Methodology in business ethics research: A review and critical assessment. *J. Bus. Ethic.* **1990**, *9*, 457–471. [[CrossRef](#)]

56. Rodenburg, K.; MacDonald, K. Enhancing Business Schools' Pedagogy on Sustainable Business Practices and Ethical Decision-Making. *Sustainability* **2021**, *13*, 5527. [CrossRef]
57. Gordon, S. Lang School of Business and Economics. Available online: <https://www.uoguelph.ca/lang/> (accessed on 12 April 2021).
58. Schmitt, N. Uses and abuses of coefficient alpha. *Psychol. Assess.* **1996**, *8*, 350–353. [CrossRef]
59. Sijtsma, K. Reliability Beyond Theory and Into Practice. *Psychometrika* **2008**, *74*, 169–173. [CrossRef]
60. Yang, Y.; Green, S.B. Coefficient Alpha: A Reliability Coefficient for the 21st Century? *J. Psychoeduc. Assess.* **2011**, *29*, 377–392. [CrossRef]
61. Guilmette, M.; Mulvihill, K.; Villemare-Krajden, R.; Barker, E.T. Past and present participation in extracurricular activities is associated with adaptive self-regulation of goals, academic success, and emotional wellbeing among university students. *Learn. Individ. Differ.* **2019**, *73*, 8–15. [CrossRef]
62. Moorhouse, T.P.; D'Cruze, N.C.; Macdonald, D.W. The effect of priming, nationality and greenwashing on preferences for wildlife tourist attractions. *Glob. Ecol. Conserv.* **2017**, *12*, 188–203. [CrossRef]
63. Howard, G.; Roe, B.E.; Nisbet, E.C.; Martin, J.F. Hypothetical Bias Mitigation Techniques in Choice Experiments: Do Cheap Talk and Honesty Priming Effects Fade with Repeated Choices? *J. Assoc. Environ. Resour. Econ.* **2017**, *4*, 543–573. [CrossRef]
64. Caillier, J. The Priming Effect of Corruption and Bureaucracy Bashing on Citizens' Perceptions of an Agency's Performance. *Public Perform. Manag. Rev.* **2018**, *41*, 201–223. [CrossRef]
65. Tomaras, P.; Tsalikis, J. Priming Effects on Cross-Cultural Business Ethical Perceptions. In *Sustainable Transport Development, Innovation and Technology*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 211–216.
66. Boer, D.; Hanke, K.; He, J. On Detecting Systematic Measurement Error in Cross-Cultural Research: A Review and Critical Reflection on Equivalence and Invariance Tests. *J. Cross-Cult. Psychol.* **2018**, *49*, 713–734. [CrossRef]
67. Boring, A. Gender biases in student evaluations of teaching. *J. Public Econ.* **2017**, *145*, 27–41. [CrossRef]
68. Gershenson, S.; Holt, S.B.; Papageorge, N.W. Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Econ. Educ. Rev.* **2016**, *52*, 209–224. [CrossRef]
69. Fan, Y.; Shepherd, L.J.; Slavich, E.; Waters, D.; Stone, M.; Abel, R.; Johnston, E.L. Gender and cultural bias in student evaluations: Why representation matters. *PLoS ONE* **2019**, *14*, e0209749. [CrossRef]
70. Pew Research Center. Political Typology Quiz. 2017. Available online: <https://www.pewresearch.org/politics/quiz/political-typology/> (accessed on 12 August 2020).