Article

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

Kathleen Rodenburg [1]*, Taimoor Rizwan [1], Ruifeng Liu[1] and Julia Christensen Hughes [1]

[1] Lang Business School, University of Guelph, Guelph 519, Canada; trizwan@uoguelph.ca; ruifeng@uoguelph.ca; jchriste@uoguelph.ca

* 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 new comer 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 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 scale could be improved. Additionally, biases in scores may exist dependent 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

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 behaviours that are at odds with the achievement of the United Nations Sustainable Development Goals (SDGs) [1,2] and 2030 agenda. In response to this recognition, a new rating system was recently developed known as the Positive Impact Rating (PIR) [3]. The PIR uniquely centers on the perceptions of students, including 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 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 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. Lang is known for its commitment to developing sustainable leaders and participated in the first PIR assessment. The efficacy of the PIR was empirically assessed on two fronts. First, a confirmatory factor analysis (CFA) was conducted to verify the factor structure of the 20- observed responses that informed the PIR scale. The CFA tested the hypothesis that a relationship exists between these observed responses and their underlying latent constructs; specifically, the separate elements of energizing, educating and engaging. Secondly, additional demographic and sociocultural information was collected in order to determine significant influences on the PIR scores, to identify potential biases that may exist when implementing the survey.

Our paper begins with a brief synopsis of two events, which focused on the shortcomings of traditional business school rankings, held in Davos, during the World Economic Forum. Following, the observations made are compared with the extant literature, with significant support being found. Next, we summarize our methods and results. Our discussion centres on the need to ensure that the measurement criteria for a chosen rating system support goals that enable a school to achieve their vision and mission. In doing so, when used effectively as a recruitment tool, it could help attract students whose goals and values are aligned, strengthening organizational culture and further supporting the school's ability to achieve its objectives. The Positive Impact Rating (PIR) holds promise to over-come the perverse effects created by traditional business school ranking systems, and therefore it is essential that the metrics used be – and perceived as - both valid and reliable. In the spirit of future enhancements to this promising scale we make 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 the17SDGs, in order to compare results based on contrasting goals and student demographics to further confirm reliability.

2. Challenges Levelled at Traditional Business School Rankings

During the 2019 World Economic Forum, 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)) (for a detailed summary

of the event see Christensen Hughes & Sommer, 2019)[4]. Here, business school deans of Principles of Responsible Management Education (PRME) Champion Schools - business schools recognized by the United Nations Global Compact as leaders with a commitment to responsible business education – alongside 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 is overemphasized; b) business schools are penalized in the rankings for turning out graduates who work for non-profits; c) course content is not evaluated; and d) teaching quality, sustainability and business ethics are minimized or absent." They concluded with a call for the re-imagination of rankings, in order 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" [4] (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, were discussion groups that considered three questions: Why do rankings matter and to whom? What is most unfortunate about current rankings? And what is our 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 incented to "play the game" and engineer results, particularly in the areas of student salaries and faculty research" [4] (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. Also, 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 [5], 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 inter-disciplinary ways.

During the 2020 World Economic Forum, 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 each of Corporate Knights, the PIR and the Financial Times; and group discussions around three key questions, which focused on: 1. participant reaction to the PIR; 2. perceptions of changes to other rankings that are underway; 3. wishes for further change [6].

Findings from the discussions suggested that there was broad support for the PIR and its focus on student perceptions, as well as its five ratings 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 (in order 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" [7] (p. 3).

3. Confirmation via the Extant Literature

While much of the focus at the events in Davos were 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 (Financial Times, 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.

3.1. Motivation

Published ranking systems originated from the need for business schools to be more forward facing toward their customers (students and businesses). According to Khurana [8], 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 an improvement in the ranking of business schools leads to a surge in the number of applicants in the coming years [9] and these rankings are more

influential than other types of media in helping potential MBA applicants determine their preferred school [10]. Elebeck [11]   found that students who graduated from highly ranked schools performed better and had higher salaries upon graduation [12]..

Another reason for rankings being increasingly valued is due to the internationalization of management education and the fact that business programs are now being marketed to potential students around the world [13]. Additionally, recruiters use these rankings to target potential employees [9]. In other research, rankings have been shown to drive resources from external partners when they too value the hierarchy of prestige in higher education [14]. However, as Pitt-Watson & Quigley [2] and others have pointed out, the metrics being used to come up with these rankings have proven troublesome.

3.2. Perverse Effects

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 [9]. For example, the Financial Times focuses 50% of their ranking criteria on the weighted salaries of alumni as well as the publication record of faculty members in selected academic journals (FT50). Few measure 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 influence both school and student behaviour, and how faculty and administrators assess the reputation and prestige of the institution [14-18].

Given the financial implications that accompany the top ranks, Athavale et al. [19] noted that some deans were at risk of losing their jobs if their published rankings fell. As a result, these measurement systems have the power to influence organizational mission, strategy, personnel and recruitment decisions, and public relation priorities [20,21].

Key findings suggested that ranking methodologies i.e., the use of subjective weights and attributes, were problematic and were open to gaming [22]. For example, a school that values rankings based on publication in certain journals puts pressure on faculty to publish in these designated 'A' journals [23]. The standard favored by 'A' category journals devalues the research published elsewhere, regardless of the content and its contribution [23]. Failure to demonstrate publishing power in these supposedly 'elite' journals risks being put on probation, being not eligible for promotion or being given a contract for teaching only [23]. Rynes [24] writes that the prestigious rankings market themselves as measuring research productivity but in essence they are only measuring 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 seem to be focused towards categorizing academics into those who are successful in garnering academic prestige and those who are not [25]. One of the impacts of this has been discouraging innovation in research. Priorities have focused on improved rankings at the expense of furthering knowledge [26-28].

Alder and Harzing [25] 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. One of the reasons their influence is so powerful is because rankings have been found to   influence hiring and promotion decisions [16,17]. Bennis & O'Toole [29] noted that rankings are so powerful that even 'rational, well-informed, well-intentioned' faculty and institutional administrators will take and support actions that focus on the interest of achieving high rankings, even when such actions undermine strategy, teaching objectives, service goals and consequently, society.

More recent contributions have further extended understanding of the perverse effects of traditional rankings and their metrics [22,30,31]. For example, Dearden, Grewal and Lilien [30] 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 research capability of business schools, which forms a heavy component of ranking metrics, introduces the risk of problematic research practices. Hall and Martin [31] 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 [31].

Johnson and Orr [32] 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 rigour and academic autonomy; as a tokenistic effort at practitioner engagement; or as a welcome development that enables scholars to embrace the pursuit of actionable knowledge" [32](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. While one business practitioner shared that they were "… staggered at the trivial nature of papers that come out and the, almost invisible veneer of knowledge that seems to be added to the world as we know it"[32](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 [33,34].

3.3. What's Next

Several universities and business schools have refused to take part and have requested not to be included in the rankings [35]. However, research has shown that non-participation could lead to organizational illegitimacy [36]. Hence many business schools, although in disagreement with the methodologies and the consequences from pursuing these ranking, still participate [36], shifting their priorities to align. Unfortunately, this shift has made business schools less adaptive to current demands by students and other stakeholders. Cutter [37] found that for the past couple of years, MBA applications have seen a sharp decline in the U.S. [38]. Technological advancements have enabled customized learning programs more aligned with student needs, but unfortunately business schools have not been leading this change [37]. Furthermore, Cutter [37] noted that business schools

tend to be viewed internally by administration as focusing on 'big money' initiatives and as 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 will 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 [39] 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 [40]. Although some have recognized their potential role in achieving the UN SDGs, many have not been engaged [41]. Edwards et al [42] 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 [43]. 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. Most business schools teach CSR and business ethics as separate standalone courses and by offering one or two courses with these assigned topics can easily obtain a full score. 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 [44]. 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 behaviour of students [45].

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 [44]. Corporate Knights [46] aligns its criteria to measure a school's contribution 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 & 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 17SDGs. 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 in 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] (p.p. 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 their goals. 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. The Study

This study is phase one of a broader study, toward uncovering embedded biases and incongruencies in methodological data collections procedures within business school ratings and rankings. Selecting the correct ranking and rating system to benchmark the organizational performance and ensuring a more valid and accurate ranking system serves to enhance institutional legitimacy by promoting behaviours internally that align with the school's vision, core values and strategy.

A controlled experiment at Gordon S. Lang School of Business and Economics at the University of Guelph, Canada was conducted to look for possible selection biases resulting from the way responses were collected for the PIR rating and how these responses may have been influenced by certain survey distribution strategies, student demographics, and socio-cultural factors. The PIR, as previous described, is designed to assess the positive impact of business schools have on society by asking students to measure several factors, including school governance, culture, programs, learning methods, student support, institution as a role model and public engagement. The PIR only surveys current students; described as the first 'by students and for students rating' [3]. The PIR as a benchmark in theory realizes the importance of business schools and their role in achieving the 17 SDGs [47] and therefore was selected as an aligned measurement tool for Lang given the Schools' PRME affiliation and its vision and strategic goals. This study was designed to answer the following research questions:

1. Does a relationship exist between the observed responses and their underlying latent constructs; specifically, do the responses to the survey support the model construct (Energizing, Engaging, Educating)?
2. Is there a selection bias in channeling the data collection for the PIR through student organizations engaged in the sustainability field? If yes, in which way does it influence the PIR results?

3. Is there a selection bias in collecting PIR data from students in courses linked to sustainability? If yes, in which way does it influence the PIR results?

4. Do demographics or socio-cultural characteristics of the student influence responses and if so, in which way do they influence the PIR results?

5. Methodology

To collect data for this study, 123 undergraduate and 33 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. Students were recruited through their club associations, and various identified classes. Clubs and classes chosen for the study were identified as either sustainability focused or not.

The questionnaire consisted of 64 questions including twenty questions that form the PIR rating, assessing how students perceive their school`s current commitment to create a positive impact, twenty-three socio-cultural and demographic questions, eleven attitude and behaviour questions to establish their sustainability attitudes and behaviour score, eight political questions to establish their political leaning score and two overall satisfaction with their academic journey questions. Two 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. Students were offered a chance to win a free coffee whether they filled out the survey or not. The electronic questionnaire took approximately 40 minutes to complete. The questionnaire is available upon request from the authors.

5.1 Dependent Variables

The overall PIR score and scores of three sub-categories of the PIR system (Energizing, Educating, and Engaging) are calculated following the original methodology in the PIR report. Specifically, the arithmetic average over 20 PIR questions for each participant is defined as the overall PIR score. Three sub scores are calculated as the arithmetic average over related PIR questions. The corresponding question numbers within each sub-category can be found in Table 6.

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 Transforming (7.4 – 8.7) business school on the scale's tiered rating system. The scores for energizing and engaging sub-categories were also transforming with scores of 7.65 and 7.45, respectively. The educating sub-category scores were positioned lower at 7.29, placing Lang on the progressing (5.9 -7.3) tier within the tiered system (see Appendix A, 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 sub-categories.

	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

5.2 Explanatory Variables

Thirteen (13) explanatory variables were constructed from the 44-question survey to test the influences of survey design, survey distribution methods as well as student demographic and, socio-cultural factors on PIR scores. A sub-set of variables were direct response categorical variables. These included: 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.

5.2.1 Political leaning index score

Based on 8 questions from a pre-existing PEW survey: Where, a lower index value would suggest more liberal leaning (left leaning), and a higher index value would suggest 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 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.

5.2.2 Attitudes toward sustainability and the environment

An index score ranging from 1 to 5 was constructed per subject based on four questions related to their attitudes toward sustainability and 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'.

5.2.3 Consumption and purchase behaviour

A binary variable was created based on a series of three questions asking students about their consumption and purchase behaviour. Zero was assigned to selected responses

if the purchasing and intended consumption behaviour did not comply with sustainability concerns.

6. Results

In total we collected 156 usable surveys. Table 2 reports the summary statistics for the political leaning score(political), attitudes toward sustainability and environment score(env_belief) and the consumption and purchase behaviour score (consum). The mean score of 0.30 indicated that participants at Lang are 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 .42, below the median of .5 indicated an intended consumption and purchase behaviour marginally away from environmentally sustainable products.

Table 2. 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 3 reports the descriptive statistics for PIR scores (Mean and standard deviation). The majority of participants identified as female (65%), under-graduates (77%), non-co-op program (63%), do not belong to a club (75%), with no faith (45%), and as 'relatively' left (liberal) political leaning (57%). Close to 49% of participants were recruited from a course with a sustainability focus.

Table 3. 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
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 participants were registered in 11 different academic programs: 5 undergraduate and 6 graduate programs, respectively. For participants from the Bachelor and Commence program, 11 sub-programs are specified. Table 4 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 4. 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 Commence and others. Panel B contains all graduate programs.

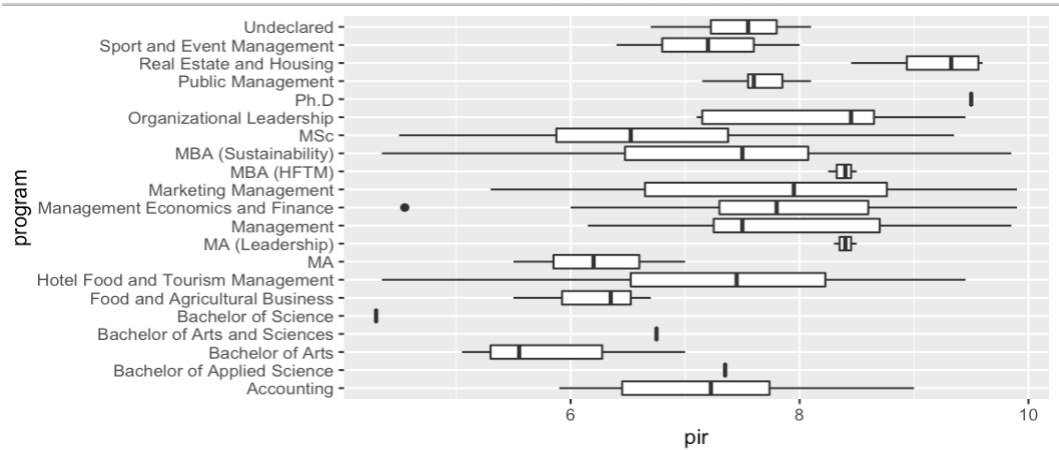| Program	Mean	N	STD
Panel A: Undergraduate ProgramsBachelor of Commence			
Accounting	7.28	16	0.95
Food and Agricultural Business	6.18	3	0.62
Hotel Food and Tourism Management	7.29	15	1.40
Management	7.84	9	1.23
Management Economics and Finance	7.75	17	1.20
Marketing Management	7.70	24	1.41
Organizational 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	7.47	4	0.59
Others			
Bachelor of Applied Science	7.35	1	NA
Bachelor of Arts	5.87	3	1.01
Bachelor of Arts and Sciences	6.75	1	NA
Bachelor of Science	4.30	1	NA
Panel B: Graduate Programs			
MA	6.23	3	0.75
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
Ph.D	9.50	1	NA

Figure 1: Box plots of PIR scores by academic discipline

There were two parts to this analysis. In part 1 a confirmatory factor analysis (CFA) was conducted to verify the factor structure of the 20- observed responses that informed the PIR scale.[1]   A CFA was chosen in this case 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 by three areas, and these were further divided into seven dimensions (See Table 5 & 6). The CFA tested the hypothesis that a relationship exists between the observed responses and their underlying latent constructs; specifically, the area of categories (Table 5) and dimensions (Table 6). 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 ≥ 0.90, TLI ≥ 0.95, and RMSEA < 0.08 would indicate a good fit.

In part 2 we conducted a bivariate (OLS) statistical model. The bivariate OLS model was a best fit[2] given the theory and the research design for this study. To understand the causes of the observed PIR score, we ran Ordinary Least Squares regressions (OLS) ensuring our data met the assumptions necessary for a linear regression to give a valid result[3]. The dependent variable in equation 1, 2, 3, 4 (Table 10) is a continuous total PIR, energizing, educating and engaging score, respectively. There are two types of explanatory variables, categorical and continuous (see variable descriptions above). Table 11 includes the same OLS analysis as table 10, but with the inclusion of interaction terms.

6.1 Confirmatory Factor Analysis

The CFA was first conducted using the latent variables (on the left) comprised of the indicators (observed variables on the right) (see Table 5 & 6). 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 5. The chi-square value(p-value) =0.00, 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.

Table 5. CFA with Three Factors. This table reports the estimated coefficients and standard errors using the Confirmatory Factor Analysis with three latent factors that are in the first column. The

---

[1] Note that the whole sample of 156 surveys are used in the CFA.

[2] The model satisfies the 7 assumptions for a linear regression, therefore providing the best estimates. Assumption 1: the dependent variable is continuous; Assumption 2: the independent variables are continuous or categorical; Assumption 3: There is a linear relationship between the dependent and independent variables; Assumption 4: There are no significant outliers; Assumption 5: There is independence of observations; Assumption 6: The data is homoscedastic; Assumption 7: the residuals (errors) of the regression line are approximately normally distributed.

[3] See footnote 2

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

Next, we conducted a CFA using the seven dimensions of: Governance, Culture, Programs, Learning Methods, Student Support, 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 Table 6 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 Seven Factors. This table reports the estimated coefficients and standard errors using the 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

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 Table 7 and 8). 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.

Table 7. 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
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

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

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

6.2 OLS Regression Analysis

OLS regressions with no interaction terms (Table 10) and with interaction terms (Table 11), with either the continuous PIR Score or three PIR sub-dimension 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 of the student that may have influenced responses.

The benchmark model is the model (1) from Table 10. The results of a multiple linear regression showed a collective significant effect of all the independent variables, $F_{(14, 128)}$ = 3.196, R^2 = 0.259. Specifically, 25.9% of the variance was explained by the model. Sustainability focused course (β= 0.625, t = 2.199, p = 0.030), Academic evaluation above (β= 0.702, t = 3.113, p = 0.003), Attitudes toward the environment (β= 0.409, t = 1.770, p = 0.080) were positive and significant in the model while, identifying with No faith (β= -0.426, t = -1.978, p = 0.051), Academic evaluation below (β= -1.072, t = -2.477, p = 0.015), Consumption behaviour (β= -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 behaviour had lower PIR scores.

To study the effect of the explanatory variables on the three sub-regions of PIR system, three more OLS regressions were run with the three PIR sub-categories scores as the dependent variables. The general effect of explanatory variables on them were similar to those on the general PIR, with several significant differences. Firstly, whether participants were from sustainability focused courses had no significant effect on the energizing dimension (β=0.383, t = 1.348, p = 0.181). On the other hand, the energizing dimension is the only one that was significantly affected by whether the participants are from co-op programs (β=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 engaging score (β= 0.789, t= 2.038, p = 0.044).

In table 11, 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 significantly lower their PIR score (β=-0.777, t = -1.744, p = 0.084). The magnitude of the influence from this interactive term on the energizing dimension (β= -0.815, t = -1.832, p = 0.070) and the engaging dimension (β=0.915, t = -1.767, p = 0.080) was similar. The effect on the educating dimension was negative but not statistically significant (β=0.699, t = -1.416, p = 0.160).

Table 10. 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	0.410	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)	(0.370)	(0.390)
Political (Right)	-0.011	-0.152	-0.018	0.219
	(0.230)	(0.230)	(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)	(0.300)	(0.317)
Constant	6.892***	7.061***	6.870***	6.695***
	(0.424)	(0.424)	(0.468)	(0.494)
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 11. 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
	(0.230)	(0.230)	(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
	(0.340)	(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**
	(0.330)	(0.330)	(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)	(0.260)
Eval (Below)	-1.110**	-0.798*	-1.379***	-0.904*
	(0.430)	(0.429)	(0.476)	(0.500)
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
	(0.270)	(0.270)	(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
R²	0.276	0.286	0.257	0.219
Adjusted R²	0.191	0.201	0.169	0.126
Residual Std. Error (df = 127)	1.187	1.185	1.314	1.380
F Statistic (df = 15; 127)	3.234***	3.387***	2.930***	2.368***

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

7. Discussion

Business school ratings and rankings serve a dual purpose. Firstly, business school rankings, signal to the community how the school is doing and serve as a powerful recruitment tool. Secondly, it is a target that influences the success rate of achieving the institution's strategic goals. The former, is driven from 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.

To this end, 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' [48]. In addition to providing a benchmark for performance by the student, 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' [48]. Building an organizational culture that is aligned with the mission, vision and core values of the institution is critical to achieve 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 validity and reliability of any chosen scale is essential. The PIR provides transparency in both the criteria used and methodologies employed. The creators are committed to develop 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 CFA analysis discovered that all questions that informed the survey are highly inter-related. Specifically, the results show a badly fitting model as the observed responses for the latent variables (i.e., energizing, educating, engaging as well as the 7 dimensions) are too interconnected and are not separate enough to clearly measure three distinct themes, or 7 separate dimensions. Table 12 & 13 highlight the high correlations between the variables.

Table 12. Correlation between 3 Latent Factors. This table reports the correlation between three sub-categories in the PIR system.

	Energizing	Educating
Educating	0.79	
Engaging	0.74	0.79

Table 13. Correlation between 7 Latent Factors. This table reports the correlation between seven sub-categories 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 due to the relatively small sample size (156). To this end, our first recommendation is to conduct a CFA analysis with a larger data set to corroborate these initial finding. 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 [49] 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 1,200. 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 when reviewing the covariance matrix.

Survey distribution methods and socio-cultural factors influenced student PIR scores. Survey distribution was not completely randomized. A subset of faculty were 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 a significant higher PIR score 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 academic experience 'exceeds expectations.' Previous research has shown that voluntary subject participation is a common phenomena of social science research and influences sampling bias [50]. Zimran [51] found that using data from non-random samples introduces endogeneity bias leading to a higher probability of affecting research that involves questionnaires. It is critically important when implementing the survey to select a random sample of students, that is representative of the entire student population. Ensuring the survey is distributed to students who do not belong to a club, or who study courses within a discipline with less environmental focus is essential for a true and honest report. If the intention is to deliver against the goals set out by the institution, then a true representation of all student experiences as a starting point is essential to identify the areas that need improvement. Interestingly, the results of this study suggest that if the published rating attracts students who are aligned with the goals of 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 club requires 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 [52]. 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.

This study also tested explicitly for priming effects. In one survey treatment, sociocultural, attitudinal, 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 behaviours of individuals [53-56].

Questions were included in the survey to identify faith affiliation, sustainable purchase and consumption behaviour 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, Ontario, Canada) students were mainly left leaning (liberal), and political orientation had no impact on PIR scores. However, those that identified with 'no faith' affiliation had a lower PIR than students who identified with faith. A student's higher environmental beliefs in terms of consumption and purchase behaviour also had a higher PIR score. Literature has shown that sociocultural attributes could lead to biased results of surveys [57-61]. Although socio cultural differences are assumed in research involving humans, the results can be interpreted wrongly if there is no comparability [57].

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. Not only does this allow the institutions to test alignment of 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 vision.

7.1 Future Research Suggestions

Future research suggestions include: 1. conducting a confirmatory factor analysis (CFA) on a larger data set to determine the latent structure equation of the survey; 2. a review of the current survey questions to ensure separability of category themes selected and provide a 'best fit model' to measure these criteria; 3. identify 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 scales ability to measure a school's positive social impact.

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

Table AA. 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

References

1. 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, U.K, 2010; pp. 117-141

2. Pitt-Watson, D.; Quigley, E. Business School Rankings for the 21st Century. UN Principles for Responsible Management Education (PRME). January 2019. Available online: https://www.unprme.org/resource-docs/ 60555MBAREPORT0119pr03.pdf (accessed on 30 March 2021).

3. Dyllick, T.; Muff, K. A Positive Impact Rating for Business Schools: Case Study. *Sustainability* **2020**, *12*, 9551, doi:10.3390/su12229551.

4. 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://grli.org/wp-content/uploads/2019/02/Davos-Dialogue-11-Feb-2019-FINAL.pdf (accessed on 28 April 2021)

5. Dyllick, T. Responsible management education for a sustainable world: The challenges for business schools. *Journal of Management Development* **2015**, *34*, 16-33, doi:10.1108/JMD-02-2013-0022

6. 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://d46c064f-971e-4554-b4f9-0441a70b5a46.usrfiles.com/ugd/d46c06_bd36328275bb4f16aec9d0ca2b7f91a0.pdf (accessed on 9 April 2021).

7. United Nations. (2015). Transforming our world: the 2030 Agenda for Sustainable Development. Available online: https://sustainabledevelopment.un.org/post2015/transformingourworld/publication (accessed on 28 April 2021)

8. Khurana, K. From Higher Aims to Hired Hands; Princeton University Press: Princeton, NJ, USA, 2007

9. Peters, K. Business school rankings: content and context. *Journal of Management Development* **2007**, *26*, 49-53, doi:10.1108/02621710710720086.

10. Tyson, G., & Greif, G. (2001). A LOOK BEHIND THE NUMBERS Editors and number crunchers explain how their b-school rankings are determined and what they do to ensure the outcome is fair. SELECTIONS-LOS ANGELES THEN MCLEAN-, 1(2), 8-15.

11. Elbeck, M. Advancing the relationship between business school ranking and student learning. *Journal of Higher Education Policy and Management* **2009**, *31*, 81-88, doi:10.1080/13600800802559294.

12. Wilkins, S.; Huisman, J. UK business school rankings over the last 30 years (1980–2010): Trends and explanations. *Higher Education* **2012**, *63*, 367-382, doi:10.1007/s10734-011-9446-7.

13. Moon, H. The globalization of professional management education, 1881--2000: Its rise, expansion, and implications. Unpublished doctoral dissertation, Stanford University, Stanford, 2002

14. Bastedo, M. N.; Bowman, N. A. US News & World Report college rankings: Modeling institutional effects on organizational reputation. *American Journal of Education* **2010**, *116*, 163- 183, doi: 10.1086/649437.

15. Elsbach,K. D.; Kramer, R. M. Members' responses to organizational identity threats: Encountering and countering the Business Week rankings. *Administrative science quarterly* **1996**, 442- 476, doi:10.2307/2393938.

16. Espeland, W. N.; Sauder, M. Rankings and reactivity: How public measures recreate social worlds. *American journal of sociology* **2007**, *113*, 1-40, doi:10.1086/517897.

17. Sauder, M.; Espeland, W. N. The discipline of rankings: Tight coupling and organizational change. *American sociological review* **2009**, *74*, 63-82, doi:10.1177/000312240907400104.

18. Sauder, M.; Fine, G. A. Arbiters, Entrepreneurs, and the Shaping of Business School Reputations. *Sociological Forum* **2008**, *23*, 699-723, doi: 10.1111/j.1573-7861.2008.00091.x.

19.   Athavale, M.; Bott, J.; Myring, M.; Richardson, L. Deans' perceptions of published rankings of business programs. *Journal of Education for Business* **2017**, *92*, 81-88, doi:10.1080/08832323.2017.1289885.

20.   Hazelkorn, E. The impact of league tables and ranking systems on higher education decision making. *Higher education management and policy* **2007**, *19*, 1-24, doi:10.1787/17269822.

21.   Hazelkorn, E. Learning to live with league tables and ranking: The experience of institutional leaders. *Higher Education Policy* **2008**, *21*, 193-215, doi:10.1057/hep.2008.1.

22.   Dasdan, A. Van Lare, E., & Zivaljevic, B. (2020). How Reliable are University Rankings?. arXiv preprint arXiv:2004.09006.

23.   Mingers, J.; Willmott, H. Taylorizing business school research: On the 'one best way' performative effects of journal ranking lists. *Human Relations* **2013**, *66*, 1051-1073, doi:10.1177/0018726712467048.

24.   Rynes, S. L.; Giluk, T. L.; Brown, K. G. The very separate worlds of academic and practitioner periodicals in human resource management: Implications for evidence-based management. *Academy of management journal* **2007**, *50*, 987-1008, doi:10.5465/amj.2007.27151939.

25.   Adler, N. J.; Harzing, A. W. When knowledge wins: Transcending the sense and nonsense of academic rankings. *Academy of Management Learning & Education* **2009**, *8*, 72-95, doi:10.5465/amle.2009.37012181.

26.   Lee, N. Reflections on assessing academic quality in marketing, and the UK REF. European *Journal of Marketing* **2011**, *45*, 477-483 doi:10.1108/03090561111111299.

27.   Macdonald, S.; Kam, J. Ring a ring o'roses: Quality journals and gamesmanship in management studies. *Journal of Management Studies* **2007**, *44*, 640-655, doi:10.1111/j.1467-6486.2007.00704.x.

28.   Macdonald, S; Kam, J. Quality journals and gamesmanship in management studies. *Management Research News* **2008**, *31*, 595-606, doi:10.1108/01409170810892154

29.   Bennis, W. G.; O'Toole, J. How business schools have lost their way. *Harvard business review* **2005**, *83*, 96-104.

30.   Dearden, J. A.; Grewal, R.; Lilien, G. L. Strategic manipulation of university rankings, the prestige effect, and student university choice. *Journal of Marketing Research* **2019**, *56*, 691-707, doi:10.1177/0022243719831258.

31.   Hall, J.; Martin, B. R. Towards a taxonomy of research misconduct: The case of business school research. *Research Policy* **2019**, *48*, 414-427, doi:10.1016/j.respol.2018.03.006.

32.   Johnson, S.; Orr, K. What is business school research for? *Academic and stakeholder perspectives, politics and relationality. Studies in Higher Education* **2020**, *45*, 557-578, doi:10.1080/03075079.2018.1564901.

33.   Skapinker, M. Why business ignores the business schools. Financial Times, January 8 2008

34.   Pfeffer, J.; Fong, C. T. The end of business schools? Less success than meets the eye. *Academy of management learning & education* **2002**, *1*, 78-95, doi:10.5465/amle.2002.7373679.

35.   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)

36.   Wedlin, L. *Ranking business schools: Forming fields, identities and boundaries in international management education*, Edward Elgar Publishing, Gloucestershire, U.K, 2006

37.   Cutter, C. Elite M.B.A. Programs Report Steep Drop in Applications. Wall Street Journal, 15 October 2019

38.   Schlegelmilch, B. B. Why business schools need radical innovations: Drivers and development trajectories. *Journal of Marketing Education* **2020**, *42*, 93-107, doi:10.1177/0273475320922285.

39.   Morgeson, F. P.; Nahrgang, J. D. Same as it ever was: Recognizing stability in the BusinessWeek rankings. *Academy of Management Learning & Education* **2008**, *7*, 26-41, doi:10.5465/amle.2008.31413860.

40.   Cornuel, E. The role of business schools in society. *Journal of Management Development* **2005,** *24*, 819-829, doi:10.1108/02621710510621321.

41.   Weybrecht, G. *The sustainable MBA: A business guide to sustainability,* 2nd ed.; John Wiley & Sons: West Sussex, U.K, 2013

42.   Edwards, M.; Brown, P.; Benn, S.; Bajada, C.; Perey, R.; Cotton, D.; Waite, K. Developing sustainability learning in business school curricula–productive boundary objects and participatory processes. *Environmental Education Research* **2020**, *26*, 253-274, doi:10.1080/13504622.2019.1696948.

43.   Storey, M.; Killian, S.; O'Regan, P. Responsible management education: Mapping the field in the context of the SDGs. *The International Journal of Management Education* **2017**, *15*, 93- 103, doi: 10.1016/j.ijme.2017.02.009.

44.   Sroufe, R. Business Schools as Living Labs: Advancing Sustainability in Management Education. *Journal of Management Education* **2020**, *44*, 726-765, doi:10.1177/1052562920951590.

45.   Hay, R., & Eagle, L. Impact of integrated sustainability content into undergraduate business education. *International Journal of Sustainability in Higher Education* **2020**, *21*,131-143 doi:10.1108/IJSHE-05-2019-0174.

46.   Corporate Knights. Available online: https://www.corporateknights.com (accessed on 4 April 2021)

47.   Mio, C.; Panfilo, S.; Blundo, B. Sustainable development goals and the strategic role of business: A systematic literature review. *Business Strategy and the Environment* **2020**, *29*, 3220-3245, doi:10.1002/bse.2568

48.   Gordon. S. Lang School of Business and Economics. Available online: https://www.uoguelph.ca/lang/ (accessed 12 April 2021)

49.   Kline, R. B. *Principles and practice of structural equation modeling*, 4th ed.; Guilford publications: New York, U.S.A, 2015

50.   Cheung, K. L.; Peter, M.; Smit, C.; de Vries, H.; Pieterse, M. E. The impact of non-response bias due to sampling in public health studies: a comparison of voluntary versus mandatory recruitment in a Dutch national survey on adolescent health. *BMC public health* **2017**, *17*, 1-10, doi:10.1186/s12889-017-4189-8.

51.   Zimran, A. Sample-selection bias and height trends in the nineteenth-century United States. *The Journal of Economic History* **2019**, *79*, 99-138, doi:10.3386/w24815.

52.   Guilmette, M.; Mulvihill, K.; Villemaire-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. *Learning and Individual Differences* **2019**, *73*, 8-15, doi: 10.1016/j.lindif.2019.04.006

53.   Moorhouse, T. P.; D'Cruze, N. C; Macdonald, D. W. The effect of priming, nationality and greenwashing on preferences for wildlife tourist attractions. *Global ecology and conservation* **2017**, *12*, 188-203, doi:10.1016/j.gecco.2017.11.007.

54.   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?. *Journal of the Association of Environmental and Resource Economists* **2017**, *4*, 543-573, doi: 10.1086/691593.

55.   Caillier, J. The priming effect of corruption and bureaucracy bashing on citizens' perceptions of an agency's performance. *Public Performance & Management Review* **2018**, *41*, 201-223, doi:10.1080/15309576.2018.1431138.

56.   Tomaras, P.; Tsalikis, J. Priming Effects on Cross-Cultural Business Ethical Perceptions. *Strategic Innovative Marketing* **2017**, 211- 216, doi: 10.1007/978-3-319-33865 1_27.

57.   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. *Journal of Cross-Cultural Psychology* **2018**, *49*, 713-734, doi:10.1177/0022022117749042.

58.   Boring, A. Gender biases in student evaluations of teaching. *Journal of public economics* **2017**, *145*, 27-41, doi:10.1016/j.jpubeco.2016.11.006.

59.   Gershenson, S.; Holt, S. B.; Papageorge, N. W. Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Economics of education review* **2016,** *52*, 209-224, doi:10.1016/j.econedurev.2016.03.002.

60.   Fisman, R.; Sarkar, A.; Skrastins, J.; Vig, V. Experience of communal conflicts and intergroup lending. *Journal of Political Economy* **2020**, *128*, 3346-3375, doi: 10.1086/708856.

61.   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, doi:10.1371/journal.pone.0209749.