1. SUMMARY

Beginning in September 2022, the Higher Education Sustainability Initiative (HESI) convened meetings of members of a Joint Task Force on Outcomes and Impacts (JTOFI). The membership of the JTOFI is drawn from two HESI Action Groups: Ratings, Rankings and Assessment (RRA) and SDG Publishers Compact Fellows (SDGPCF).

The Task Force was asked to explore how higher education stakeholders could move towards a more outcome focused way of interpreting their work on sustainability in general and the SDGs in particular. At the outset it was recognised that academic assessment is often focused on research, but the role of higher education in sustainability is considerably wider, incorporating teaching, operations, and outreach, and these aspects are often under-assessed.

Considerable work by governments, funders, higher education institutions and RRA organizations aimed at achieving a more outcomes-based focus for ratings, rankings and assessments is already underway. The United Nations has implemented the Principles for Responsible Management Education (PRME) initiative, for example, and its work with the Globally Responsible Leadership Initiative (GRLI) is also highly relevant.

Furthermore, this work is not meant to repeat or replace the previous work of the RRA group or the SDGPCF. Instead, it tries to build on existing recommendations. It is useful to repeat the definition of Rankings, Ratings and Assessment organisations from HESI assessment for the SDGs - Volume 1 Assessors.

“...organisations and individuals who are involved in the creation and implementation of higher education assessments that use, or incorporate elements of, sustainability and especially the UN SDGs. This explicitly includes ratings and rankings. .... Organizations involved in this work may be private companies, universities, or governments.”
Taking note of emerging international trends (see below), this report on the work of the JTFOI recommends the following next steps:

**Enhancing measurement**

- **Governments** to develop assessment frameworks that take the real-world impact of research linked to sustainable development into account.
- **Governments and national higher education associations** to acknowledge and support Voluntary University Reviews.¹
- **Higher education institutions (HEIs)** to adopt alternative research assessment tools for internal and external evaluations that focus on outcomes.
- **HEIs** to engage with and provide feedback to RRA organizations demanding that outdated systems of measurement are no longer used, and to highlight the work of those RRA organizations that develop improved systems.
- **RRA organizations** to include more research published outside scholarly journals in their systems of measurement, as it becomes available.

**Practice**

- **Governments and national higher education associations** worldwide to make national level commitments to the role of higher education in delivering sustainable development, and to commit to assessing the outcomes of this work.
- **Publishers** to commit to commissioning and publishing work that connects academic researchers with practitioners and policymakers.
- **HEIs** to survey more groups of stakeholders when compiling the data they provide to RRA organizations, including students and practitioners. **RRA organizations** should consider the possibility of including these groups in their activities.
- **HEIs and RRA organizations** to support open science as defined by UNESCO and incorporate openness as a data value.
- **HEIs and RRA organizations** to advocate for and implement transparency in rankings criteria.

**Funding**

- **All funders of higher education, including private foundations** which fund research, to prioritize research which has a real-world impact linked to sustainable development.
- **All funders of higher education** to commit to appropriate funding of the measurement of outcomes as part of their research funding.

**General**

- Where appropriate, tier by type of HEIs (e.g., research intensive institutions, teaching-focused institutions, etc.) and within countries/regions. Also provide flexibility to allow for regional outcomes, not just global outcomes.
- Assessments of research and education should include qualitative as well as quantitative elements, the former providing context for the latter.

¹ See Universities Canada (2023), 10.
• “Applied research” should carry the same weight as “pure research” for RRA organizations, universities, and other stakeholders, and not be regarded as inferior when measured in fulfillment of the SDG criteria.

• Research and education initiatives that are explicitly co-designed and co-produced with intended beneficiaries and/or practitioners should be recognized and rewarded in systems of measurement including ranking systems.

• All stakeholders should look for ways to recognize and reward published academic outputs that contain implications for policymakers in an Abstract, or that contain a Plain Language Summary (PLS) and/or Plain Language Summary of Publication (PLSP).²

• All stakeholders should consider ways to evaluate and report on students’ knowledge, skills and values longitudinally, i.e., by measuring levels of competence pre- and post- efforts towards Education for Sustainable Development (ESD), as well as shifts in values, including intentions to act more sustainably and pursue sustainable/purpose-focused careers.

• Assessments need not be limited to research and education but may also take operational considerations into account. Green campus initiatives such as recycling initiatives, reduced waste, reduced electricity use, more efficient water use, etc., produce measurable outcomes.

2. INTRODUCTION

Beginning in September 2022, the Higher Education Sustainability Initiative (HESI) convened meetings of members of a Joint Task Force on Outcomes and Impacts (JTFOI). The membership of the JTFOI is drawn from two HESI Action Groups: Ratings, Rankings and Assessment (RRA) and SDG Publishers Compact Fellows.

Attendees (1 or more meetings) included: Mohamed Aisati, Gerald Beasley (co-chair), Jean-Christophe Carteron, Julia Christensen Hughes, Thomas Dyllick, Nikolay Ivanov, Andrew Jack, Nicola Jones, Leigh Kamolins, Nikita Lad, Ratna Lubis, Drew MacFarlane, Carina Mutschele, Kathleen Ng (co-chair), John North, Angela Partridge, Jonghwi Park, Jay Patel, Tony Roche, Duncan Ross (co-chair), Debra Rowe, Joshua Sowah, Lucas Toutloff, Roger Worthington.

The JTFOI has met “virtually” 8 times for 60 or 90 minutes to date: September 1 & 29, 2022; October 20, 2022; Dec. 1, 2022; Jan. 5, 2023; Feb. 17, 2023; March 20, 2023; May 15, 2023.

The Joint Task Force on Outcomes and Impacts (JTFOI) was established to consider one of the most important – and challenging - issues raised in vol. 1 of the HESI Rankings, Ratings and Assessment Action Group’s Assessments of...

Higher Education’s Progress towards the UN Sustainable Development Goals, i.e., how to measure real change in the world when assessing the contributions of higher education to sustainable development. Monthly meetings of the Task Force have brought together a small group of stakeholders, all of whom wish to see higher education fulfil its potential to advance the world towards achieving the Sustainable Development Goals. One important way to encourage higher education to do this would be to develop and promote ways that RRA organizations could usefully measure real world outcomes. The Task Force was accordingly established to provide guidance to RRA organizations seeking to incorporate meaningful assessments of sustainable development into their criteria. It was also an opportunity for high-level participants to try out ideas by expressing them to a diverse but sympathetic group of people.

In the initial meetings co-chair Kathleen Ng oriented the group with presentations drawn from management theory on managing complexity (as opposed to managing simple or complicated tasks); the roles of appreciative inquiry; the strength theory of change; and the value of establishing minimum specifications.

The Task Force has also remained cognisant of landscape changes that have taken place even since its formation last September, notably the increased frequency of climate change-related disasters in the so-called Global North; the uptick in negative media coverage of university rankings; and the Nov. 2022 research release of OpenAI’s ChatGPT.

This Report is submitted to document the work to date of the Task Force for HESI and other interested parties. It includes recommendations for RRA organizations seeking to incorporate sustainable development into their criteria as well as next steps towards identifying further recommendations.

3. DEFINITIONS

RRA systems to rate, rank or assess HEIs according to their contributions to the SDGs currently rely on HEI outputs. However, it has become clear that it is more important in an SDG context to rate, rank or assess HEIs in terms of their outcomes. Definitions of inputs, outputs, outcomes and impacts are varied and typically depend on context and purpose. The context and purpose in this case is global sustainable development. Working definitions used in this report are as follows:

- Inputs: resources (which produce outputs)
- Outputs: products (which support outcomes)
- Outcomes: effects (which support impacts)
- Impacts: long-term effects

So the appropriate deployment of resources (inputs) will produce products (outputs) that support the SDGs (outcomes) and lead to sustainable development (impact).
These four concepts form a chain. Outputs and outcomes are critically important links in the chain because they connect inputs to impacts. At each stage, the links may be visible or invisible, intended or unintended, equitable or inequitable, etc.

In our context, scholarly publications and other products of HEI are best regarded as outputs. No scholarly publication can be regarded as itself an outcome or impact because its existence does not in itself contribute to sustainable development no matter what its quality, format, etc. Importantly, this is also true of educational initiatives by HEIs such as public lectures, public access to computers, etc.

It is however true that each scholarly publication supports one or more outcomes. Many of these outcomes are well-known and well-researched, e.g. advancing knowledge; gaining tenure or promotion; enhancing institutional reputation. They have been measured for decades if not centuries.

Similarly, student retention and ultimately, the conferring of degrees is an outcome, following demonstrated competence (an output) within a particular program of study. The curriculum, learning activities and assessments, alongside the pedagogical skill and commitment of the faculty are inputs. The impact is found in the behaviour and career path of the student.

Unfortunately, but not surprisingly, none of these outcomes – and none of the surrogates measured to assess them – has anything to do with sustainable development in practice. It is not really worth measuring how many SDG-related research articles are published by an HEI unless the sustainable development outcome of the articles can also be determined. Similarly, it is not really worth counting how many computers in the library are publicly accessible if no member of the public uses them and changes their behaviour as a result. If all we do is advance our knowledge of poverty, for example, we have not even begun to address SDG 1. And if we offer a course on sustainability, but few students take it, or if little is remembered at its conclusion, the impact on future behaviour impact is likely to be minimal.

In these circumstances, measuring outputs as indicators of outcomes is risky. It is still important to make it easier to bridge the gap between researcher and practitioner at the article level by asking authors to identify the most relevant SDG, itemizing practical applications of their research using Plain Language Summaries, etc. But it will be hard to assess the value of all this good work unless HEI outputs can be connected to sustainable development outcomes.

4. TYPES OF MEASUREMENT

In addition to non-affiliated stakeholders, the JTFOI welcomed the participation of representatives from a variety of organizations that use different types of rating, ranking, or assessment. Participants included representatives of organizations that took responsibility for universities (e.g. Times Higher Education; QS; Sulitest), business schools (e.g. Financial Times, Positive Impact Rating), academic journals (e.g. Elsevier, Emerald Publishing) and individual faculty
members (e.g. business schools, universities). Sustainable development - and in particular the SDGs - could be promoted via any or all of these types of measurement. However, JTFOI members recognized that each has specific issues to address when incorporating sustainable development priorities into their systems of measurement.

These issues often revolve around the difficulty of aligning sustainable development priorities with the different purposes of the measurement systems used in each assessment. In the case of individual faculty members, for example, the main purpose of the measurement system is usually to help decide on awards of promotion and tenure. This measurement system is only rarely linked to the social benefit of the faculty member’s research or teaching. In any case, there can be a considerable time delay between when research is disseminated and when its impact might begin to be felt. It would take more time than most Tenure Committees possess to add the qualitative assessments required. The JTFOI heard about the emerging trend – especially among business schools – to build progress on sustainable development into tenure and promotion evaluation criteria. But participants also recognised the difficulty of ensuring continuity given that such a change requires leadership committed to such a transformation.

5. CURRENT GAPS AND WEAKNESSES

Current systems of rating, ranking, or assessing in higher education environments often draw from data about inputs and outputs rather than outcomes. This, of course, is not unique to ratings or rankings - it is a common failure in measurement systems. So, for example, organizations tend to count how many research articles have been published on a particular topic, or how many students have graduated from a relevant course, rather than how either of these outputs have led to actual behavioral change in the target audience.

This becomes a weakness when RRA organizations seek to incorporate sustainable development into their criteria. RRA organizations are aware of this, and would like to have more outcome focused measures, but these are not easy to establish. The UN’s SDGs emphasise the importance of the practical applications of research and education. But how is it possible to measure higher education’s impact on the practice of sustainable development, and what (if any) should be the time limits applied? JTFOI participants pointed out that “pure” research could be at least as valuable as “applied” research – and that perhaps the distinction was false or at least blurry in any case. Also, research leading to changes in policy (e.g. law), even when policy is defined as an intermediate step between output and outcome, may be valuable and worth recognizing.

JTFOI participants also noted that RRA organizations are naturally predisposed to resist change. Just about all RRA organizations want to avoid continual shifts in what or how they measure. Tracking year-over-year changes (who moved up, who moved down) is a popular institutional pastime that gets compromised whenever metrics change. A constantly shifting focus serves no one. And RRA organizations have developed systems that can handle particular types of data better than others. Of course, those that rely on data provided by external
sources have also become reliant on the capacity and willingness of those sources to supply that data.

Furthermore, many of the current methods of ranking, etc., were originally developed without sustainable development in mind – though exceptions represented in the JTFOI meetings included STARS (Sustainability Tracking Assessment & Rating System); the THE Impact Rankings; the Sulitest; and the Positive Impact Rating. Task Force members recognized that the longitudinal drag felt by some systems of measurement – whereby changing what is measured makes time-based comparisons more difficult – significantly reduces the impetus for some RRA organizations to adapt their measures to new sustainable development priorities. This is especially true because sustainable development itself is an evolving field, and the priorities outlined by the Sustainable Development Goals in the UN’s 2030 Agenda will need updating before the end of the decade.

Lastly, measuring outcomes rather than outputs is always more likely to produce unwelcome volatility in any ranking system. This is because statistical comparisons that measure outcomes tend to have only limited data from which to draw conclusions - and greater freedom in how that data can be used. Furthermore, many of the factors contributing to a research or educational outcome are beyond the view or control of the person(s) or thing being measured. Outcomes are not only difficult to calculate, they are also difficult to predict. It is really not possible to expect universities, journal editors, publishers or faculty members to self-report accurately on their outcomes rather than their outputs, especially given the challenges already seen to be associated with the accuracy of self-reported data. Furthermore, measuring is always an expensive affair, and consequently what gets measured is most often what is easily counted.

6. EMERGING INTERNATIONAL TRENDS

JTFOI participants identified several emerging international trends that mitigate the gaps and weaknesses of the current systems of ranking, rating and assessing in higher education. In the list given below, each trend is linked to a recommended next step (see Summary above). An example and a caution is also given when it seems appropriate.

Enhancing measurement

1. Research assessment by governments increasingly takes the so-called real-world impact of research linked to sustainable development into account.

Example: The UK’s Research Excellence Framework (https://www.ref.ac.uk/).

Next Step: Governments to develop frameworks that take the real-world impact of research linked to sustainable development into account.

Caution: the UK’s Research Excellence Framework is extremely costly (estimated at over £400m).
2. Voluntary University Reviews (VURs) have emerged in the last few years as one important way for higher education institutions to measure and report on their contribution to the SDGs.


Next step: **Governments and national higher education associations** to acknowledge and support Voluntary University Reviews.

Caution: How to ensure validity/reliability of self-reported data, as institutions have a self-interest in being positively assessed.

3. Alternative research assessment tools are being developed to assess the relatedness of published business school research to the SDGs.

Example: The Responsible Research Assessment tool.³

Next step: **Higher education institutions (HEIs)** to adopt alternative research assessment tools for internal and external evaluations that focus on outcomes.

Caution: the time cost of assessment tools should be carefully evaluated.

4. There is an increasingly widespread dissatisfaction with current systems of research assessment (as various North American media outlets have reported, some major universities are even refusing to participate in certain rankings).

Next step: **HEIs** to engage with and provide feedback to RRA organizations demanding that outdated systems of measurement are no longer used, and to highlight the work of those RRA organizations that develop improved systems.

5. High quality research on sustainable development that includes recommendations for research outcomes is often published outside the scholarly journals indexed by databases such as Clarivate’s Web of Science or Elsevier’s Scopus. Efforts are underway to incorporate so-called grey literature such as government reports, policy documents and working papers in these databases. There are also new organizations devoted to indexing this research. It is therefore increasingly feasible to include the publication of such literature in systems of measurement.

Example: Policy Commons selects “high quality, data-rich reports from more than 24,000 policy organizations, including small think tanks and municipalities missed by others” (https://policycommons.net).

Next Step: **RRA organizations** to include more research published outside scholarly journals in their systems of measurement, as it becomes available.

---

³ See Rodenburg, K; De Silva, V; Christensen-Hughes, J. (2021).
Cautions: Deduplicating multiple datasets can lead to errors.

Practice

6. There has recently been a much-needed increase in the participation of higher education institutions worldwide in the measurement of their contributions to sustainable development.

Example: THE Impact Rankings and AASHE STARS rating combined are already reporting on the sustainable development work of some 2-2,500 higher education institutions.

Next step: Governments and national higher education associations worldwide to make national level commitments to the role of higher education in delivering sustainable development, and to commit to assessing the outcomes of this work (see example from Universities Canada https://www.univcan.ca/sustainable-development-goals/).

7. Publishers increasingly recognize the importance of impact to authors, higher education institutions and research funders.


Next step: Publishers to commit to commissioning and publishing work that connects academic researchers with practitioners and policymakers.

NB. There is a growing number of ways to achieve this next step – by, for example, convening and/or offering support to summits that match research projects with community needs; by including Plain Language Summaries (PLS) or other forms of practitioner abstracts in published research articles; by requesting that pre-publication peer reviews include an assessment of the practical value of a submitted article; by adding people with practical experience in the field to journal advisory boards.

Example: Guidance for research authors in preparing abstracts that encompass research limitations/implications, practical and social implications are provided by Emerald Publishing, for journals it publishes: How to: Write an article abstract | Emerald Publishing (emeraldgrouppublishing.com)

8. There is an emerging international trend to engage more stakeholders in the data that institutions provide to RRA organizations.

Example: Positive Impact Rating for Business Schools (https://www.positiveimpactrating.org/) engages students in rating their perceptions of the positive societal impact of the schools they attend.

Next step: HEIs to survey more groups of stakeholders when assessing their impact, including students, employers and practitioners. RRA organizations should consider the possibility of including these groups in their activities.
9. There is a movement towards a global open science environment, in which scientific research and its supporting data is equitably and openly shared. This emerging international trend towards open science aims to eliminate the financial and other barriers experienced by academic and non-academic communities seeking to access scientific knowledge.


Next step: HEIs and RRA organizations to support open science as defined by UNESCO and incorporate openness as a data value.

10. There is increasing transparency in what the ranking represents, based on the metrics used in determining the ranking.


The ranking is clearly stated to be an assessment of the ROI of an MBA, comparing tuition with earnings lift.

Next Step: HEIs and RRA organizations to advocate for and implement transparency in rankings criteria.

Funding

11. Private foundations which fund research are increasingly requiring the research they fund to have a real-world impact.

Example: The Bill and Melinda Gates Foundation, which awarded $7 billion total charitable support in 2022.

Next Step: All funders of higher education, including private foundations which fund research, to prioritize research which has a real-world impact linked to sustainable development.

12. There is an increasingly widespread recognition that measuring costs money.

Next step: All funders of higher education, including private foundations, to commit to appropriate funding of the measurement of outcomes as part of their research funding.

Caution: if funding for measurement is not provided, but measurement is expected, then this will put research outside of the reach of many poorer institutions.

General

NB. As well as responding to emerging international trends, JTFOI members adopted the following general recommendations derived from a Framework for
the Assessment of Sustainable Outcomes (FASO) that was submitted and discussed.

13. Where appropriate, RRA organizations, universities, and other stakeholders should tier by type of HEI (e.g. research intensive institutions) and within countries/regions. Also provide flexibility to allow for regional outcomes, not just global outcomes.

14. Assessments of research and education should include qualitative as well as quantitative elements, the former providing context for the latter.

15. “Applied research” should carry the same weight as “pure research” for RRA organizations, universities, and other stakeholders, and not be regarded as inferior when measured in fulfillment of the SDG criteria.

16. Research and education initiatives that are explicitly co-designed and co-produced with intended beneficiaries and/or practitioners should be recognized and rewarded in systems of measurement.

17. All stakeholders should look for ways to recognize and reward academic journal articles, etc., that contain implications for policymakers in an Abstract, or that contain a Plain Language Summary (PLS) and/or Plain Language Summary of Publication (PLSP).4

18. All stakeholders should consider ways to evaluate and report on students’ knowledge and behaviors longitudinally, i.e., by measuring levels of knowledge pre- and post- efforts towards Education for Sustainable Development (ESD) and including intentions to act more sustainably.

19. Assessments need not be limited to research and education but may also take operational considerations into account. Green campus initiatives such as recycling initiatives, reduced waste, reduced electricity use, more efficient water use, etc., produce measurable outcomes.

7. ARTIFICIAL INTELLIGENCE: OPPORTUNITIES AND LIMITATIONS

Artificial intelligence (AI) enables computer systems to perform tasks requiring human-level intelligence, such as learning, reasoning and acting based on data patterns. The rise of big data, advanced algorithms and increased computing power is driving rapid global adoption of AI across sectors. While AI innovations promise solutions to complex challenges, ethical concerns remain around transparency, bias, privacy and effects on labor. International cooperation and

governance will be critical to ensure AI benefits society as these technologies proliferate.

AI has existed as a concept since the 1940s when the term "artificial intelligence" was coined by John McCarthy, Marvin Minsky, Nathaniel Rochester and Claude E. Shannon. One of the more famous “founding fathers” of AI is Alan Turing, he is most well-known for the Turing Test, which is a way to figure out if a machine can act in a human-like way. It was not until recently that AI became a household name and a part of everyday conversations thanks to the introduction of Large Language Models (LLMs) such as GPT, Bard, Galactica, Bloom, LaMDA, PaLM, and Claude.

Large language models are a type of artificial intelligence. They are trained on massive amounts of text data. This allows them to generate and understand language in very human-like ways. Big tech companies use these models to power conversational apps like chatbots, translators, and voice assistants.

Opportunities:

Here are some ways that integrating large language models into higher education practices can help achieve Sustainable Development Goals:

- Personalized learning systems powered by LLMs expand access to quality education by adapting to each student's needs.
- Chatbots can be used to help students during enrollment and make aid processes more efficient and affordable
- Make learning more inclusive and accessible through real-time lecture translations
- Intelligent writing assistants will help students and researchers boost scholarly output, improve the quality of writing, reduce time spent searching and reading, and eliminate citation of retracted papers
- Tutors powered by LLMs provide basic skills education in remote areas lacking teachers
- Analyzing student needs data with LLMs guides the development of equitable education policies.

It is important that LLMs are integrated into existing processes in a thoughtful manner and that all stakeholders are aware of when LLMs are being used and for what purpose. We also need to be aware that LLMs are not magic and come with risks, biases, and potential harms.

Limitations:
As with all tools, LLMs have limitations and risks, here are a few that we all need to pay close attention to when implementing these solutions to help improve higher education and attain the SDGs:

- Can learn biased views if the training data has biases. This could lead to unequal treatment.
- Struggle to understand local contexts and cultures. One approach won't work everywhere.
- Black book - It's hard to explain the systems' mistakes and fix them. Inner workings are complex.
- Using student data raises concerns about privacy, consent, and monitoring.
- LLMs may lead to job displacement in education. Utilize it to make your people better, not replace them.
- Require lots of data, computing power, and storage.
- Copyright violation is a big concern and there are a growing number of lawsuits from authors, actors, musicians, and creators against the developers.
- LLMs may hallucinate nonexistent references due to lacking real world knowledge.

8. SELECT TOOLS AND RESOURCES

Publications:


Rodenburg, K; De Silva, V; Christensen-Hughes, J. (2021). SDGs: A Responsible Assessment Tool toward Impactful Business Research. Sustainability 2021, 13, 14019. [online] Available at: https://ideas.repec.org/a/gam/jsusta/v13y2021i24p14019-d706027.html


towards the UN Sustainable Development Goals. Vols. 1-3. 

Woolston, Chris (2023). How to measure the societal impact of science. Springer 

Websites:

EVALSDGs  https://evalsdgs.org/

Network For Business Sustainability  https://nbs.net/

Principles for Responsible Management  https://www.unprme.org/