

EXPLORING RESEARCHERS' VIEWS ON METRICS AND RESEARCH IMPACT

INTERNAL REPORT

SEPTEMBER 2023



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

PROJECT TEAM

Ammon Apiata, Māori and Indigenous Research Support Advisor

Fiona McCallum, Research Information Analyst

Garth Smith, Open Research Librarian

Jess Howie, Open Research Manager

Kalehi Apete, Kauvaiora

Marissa Cassin, Open Research Librarian

Mel Chivers, Open Research Librarian (Project Lead)

Yilan Chen, Open Research Librarian

ACKNOWLEDGEMENTS

The project team would like to thank and acknowledge all of the participants for their time and openness discussing this emotive and important topic.

We would also like to acknowledge the leadership from Bryony James and Michelle Blake for recognising the value of this research and providing our project team the opportunity to carry out this research.

TABLE OF CONTENTS

Acknowledgements	2
Executive summary.....	4
Introduction	5
Literature review	5
Research design	9
Research questions.....	9
Approach.....	9
Participants	9
Data analysis	10
Ethical considerations.....	10
Limitations	10
Themes.....	11
Theme 1: No one-size fits all: Impact and excellence varies greatly across disciplines and matter more to some disciplines than others	11
Theme 2: Lack of understanding of metrics and reliance on H-index.....	13
Theme 3: Metrics are subjective and can be gamed	15
Theme 4: A general nervousness around metrics or disdain, especially for Altmetrics.....	16
Theme 5: Cultural considerations and the importance of the collective	17
Theme 6: Disconnect between what they value and what they think their peers value.....	20
Theme 7: Publishing decisions are driven by a diverse range of factors.....	20
Theme 8: Varying levels of importance in getting research out to the public including OA.....	22
Conclusion.....	25
References	27

EXECUTIVE SUMMARY

Research excellence is a goal that universities and researchers alike strive for, but how do we identify and measure research excellence? Do quantitative metrics demonstrate the impact that researchers are aiming for?

This report details the results of an exploratory case study investigating how researchers responded to a range of quantitative metrics. As a part of their approach, the project team interviewed thirteen academic staff at the University of Waikato across a range of disciplines.

A thematic analysis of the interviews was carried out, bringing to light eight dominant themes from the participants' discussions:

- No one-size fits all - impact and excellence varies greatly across disciplines and matter more to some disciplines than others
- Lack of understanding of metrics and reliance on h-index
- Metrics are subjective and can be gamed
- A general nervousness around metrics or disdain, especially for Altmetrics
- Cultural considerations and the importance of the collective
- Disconnect between what they value and what they think their peers value
- Publishing decisions are driven by a diverse range of factors
- Varying levels of importance in getting research out into the public including OA

This paper argues that metrics can be used as a tool for researchers, but should be understood in the context of non-quantitative measures. Metrics alone cannot determine how impactful a researcher's contribution may be. It is important for researchers to be recognised as individuals in order for them to tell the story of their work. This will require upskilling both for researchers and those who evaluate research.

INTRODUCTION

Researcher profiles and the metrics associated with them have become ubiquitous, but the extent to which they fairly represent researchers and their work varies dramatically across disciplines, methodologies, and cultural contexts. Traditional metrics primarily measure an article's scholarly impact by tracking how frequently it is cited by other articles, using tools like the h-index to gauge the author's publication volume and citation count, and journal-level metrics, which measure citation impact for journals rather than the research itself (Cooper, 2015). As the academic research environment moves from print to digital formats and becomes more accessible online, alternative metrics, called altmetrics, have emerged that measure scholarly impact by analysing online interactions with an author's work (Bakker et al., 2020; Bornmann et al., 2019). This has led to a shift in how we view, discover, and evaluate researchers and their work, emphasising the need for a nuanced approach to dissemination methods. It has long been recognised that some disciplines are better suited to traditional metrics (Buckle & Creedy, 2019; Hicks et al., 2015). Altmetrics and a range of newer metrics normalised by field attempt to resolve these disparities, but difficulties remain (Karanatsiou et al., 2017).

Through a case study approach, this research explores the variety of profiles and metrics offered and utilised by the University of Waikato across various disciplines. It will identify tools that can help researchers better understand and communicate their research impact, particularly in disciplines that are not well-served by traditional metrics. Ultimately, this research aims to ascertain which tools are most useful for researchers who wish to understand their impact and effectively communicate their findings.

LITERATURE REVIEW

The debate around the use of metrics for assessing research excellence can be traced back to Charles Babbage in 1830, continuing through the early 2000s with the introduction of Web of Science and Google Scholar to the present day (Csiszar, 2017, Carpenter et al., 2014). This timeline reflects changing attitudes over time and in response to the evolving academic ecosystem. Today, traditional metrics are often used to demonstrate research excellence through the number of citations and publications an author has; these might include the author's h-index or the 'Journal Impact Factor' metric (JIF). Using JIF has been problematic from the start because it has mistakenly been used to evaluate the research excellence of researchers despite being designed to specifically evaluate journals (Magnus, 2013; Ali, 2021). The use of JIF has decreased in recent years, though not entirely abandoned, according to Curry et al. (2022), as the JIF is still used in conjunction with other metrics (p. 53). There is an ongoing discussion about the use and fairness of these metrics, which has led to proposals for switching to other measures of excellence (Hobson & Hall, 2010; Karanatsiou et al., 2017; Kondakci et al., 2021; Tilbury et al., 2022; Olive et al., 2022).

While the bulk of the literature is concerned with proposals for switching to other metrics, there is comparatively little in the way of endorsements to do away completely with metrics. The uptake of “metric culture” varies across disciplines and countries (Hammarfelt & Haddow 2018), with some institutions and countries working to decrease their reliance on metrics to demonstrate research impact (HuMetricsHSS, n.d.; Hicks et al., 2015; Curry et al., 2022; Price, 2022). This is in conjunction with the San Francisco Declaration on Research Assessment (DORA) which recommends the use of more metrics rather than none, albeit in a bid to distance evaluation of research excellence from using JIF alone. Alongside DORA is the Leiden Manifesto, the Metric Tide report and others that do not call for the end of metric use, but rather for the responsible use of metrics (HuMetricsHSS, n.d.; Hicks et al., 2015; Curry et al., 2022; Price, 2022).

Current discussion in the literature around metrics identifies issues such as metrics manipulation, the radically different citation patterns across disciplines, and their inappropriate use in performance research assessment (Asaolu et al., 2022; Chatterjee et al., 2020; Sarpong, 2021). Bornmann (2017) explains that because of the way scientific progress is made, by focusing on citation counts and articles published, traditional metrics do not accurately reflect a researcher’s contribution. Traditional metrics can be skewed towards the hard sciences, where citations are often higher due to greater rates of co-authorship and citation practices. Current literature suggests that traditional metrics do not properly evaluate research excellence, even in the hard sciences (Bornmann, 2017). This is further complicated because in disciplines outside of the hard sciences, dissemination can happen in a vastly different manner such as performances or industry reports, and citation of such research would also take a different form, even within the same discipline (Phillips, 2020, Tilbury et al., 2022).

In answer to this, alternative metrics (altmetrics) examine the use and attention that research receives beyond citations. Altmetrics capture social media, news and other engagement, and categorises them based on views, downloads, discussions, recommendations or endorsements, and policy citations (Karanatsiou et al., 2017, p. 21). There are, however, also limitations to using altmetrics. Studies have shown that altmetrics, including Altmetrics.com, have a bias towards the English language and North America (Yang et al., 2021). An expanding area of altmetrics captures policy citations and a number of sources (including Overton, SciVal, PlumX, and Dimensions) are now capturing citations in policy documents from non-profit (NGOs) and government organisations.

While the purpose of all these metrics is to evaluate research, Chatterjee et al. (2020) note a substantial divide between those that promote the positive aspects of research assessment and those critical of current research assessment practices. This was evidenced when comparing researchers opinions of the United Kingdom’s Research Excellence Framework (REF) and New Zealand’s Performance-Based Research Fund (PBRF) (Chatterjee et al., 2020). A “growing self-interestedness in academia” was seen as a consequence of the PBRF (Chatterjee et al., 2020, p. 1241). The concern with the rise in self-interest is that it would hamper collaboration and information sharing (p. 1241). Chatterjee et al. (2020) determined that there was a danger with the individual focus of the PBRF, suggesting that it could cause rifts within institutions (p. 1241).

This links to another concern among researchers who are concerned about the potential of gaming metrics. Graf et al. (2019) identified a list of negative behaviour markers considered to be gamification, including: breaking up a research project into several different articles to generate more

publications; topic selection based on popularity rather than relevance; questionable methodology use; and deprioritising other work (p. 761-762). Gamification of research excellence is not limited to individuals, as institutions have also been gaming the PBRF (Chatterjee et al., 2020). Buckle & Creedy (2019) detailed how performance assessment indicators such as the Average Quality Score (AQS) in PBRF assessment can negatively act as an incentive within institutional hiring practices (p. 7-9). The PBRF Review Panel (Salesa, 2020; Smith et al., 2020) acknowledged a submission for the removal of the AQS because "it is not meaningful and encourages gaming" (p. 104) and responded by recommending the indicator be discontinued (see recommendation 33).

Another concern is the underrepresentation of marginalised people across all disciplines, both nationally and globally (Chatterjee et al., 2020; Ferrier-Watson, 2019; Kondakci et al., 2021; Onyancha, 2022; Sarpong, 2021). Over the past two decades, several Māori academics have argued that their research impact is not always fully accounted for through the PBRF system (Kidman et al., 2015; Roa et al., 2009; McAllister et al., 2020; Tawhai et al., 2004). The PBRF's use of binary measures of 'research excellence' means that Māori researchers often forego research outputs and opportunities that will benefit their PBRF scores, in order to give precedence to their whānau and community responsibilities (Kidman et al., 2015, p. 81; Roa et al., 2009). This is also experienced by Pacific researchers, where traditional metrics fail to capture the quality and impact of Pacific research (Naepi, 2021). Central to Indigenous research is long-term community engagement and a focus on community needs that are not necessarily met by publishing in high-impact journals and as such are not always captured in traditional metrics; the Sāmoan axiom *teu le vā* (strengthen the relationship), is a methodological example of how Sāmoan Indigenous research is regarded and received by Pacific Indigenous researchers and their respective communities (Anae, 2016). This relational concept pertains to the inter and intrapersonal connections that are fundamentally entwined within Pacific research and assessment (Anae, 2016). This radically contrasts with the analysis of the PBRF as a neoliberal and managerialist system as discussed by Cupples & Pawson (2012). While the PBRF system's emphasis on research activity has generally been praised among New Zealand researchers, McAllister et al. (2020) and Naepi (2021) also point out that Māori and other Pacific researchers' wider contribution to the work of knowledge production is rendered invisible by the quantitative nature of traditional metrics.

Māori, Pacific, and female researchers are underrepresented in the academic workforce (McAllister et al., 2019; Smith et al., 2020). The PBRF report by Smith et al. (2020) stated that "600 female staff, almost 750 more Māori staff and around 450 more Pacific staff" would be needed if these groups were to represent their share of the national population in a funded Quality Category (p.32). The panel described these disparities as complex and influenced by implicit and unconscious biases, which was exacerbated by the lack of systemic action (Smith et al., 2020). The report's recommendations to address these discrepancies later received some critique from Naepi (2021), who argued the changes were examples of performative action and hypothesised an ensuing capitalising of Pacific bodies through financial incentives directed at the institutional level. Such marketisation of academics is also discussed at the broader level by Sarpong (2021), who explores researcher autonomy in relation to maximisation of a performance-based system.

As we consider the use of quantitative measures in both representing and determining research excellence, quality, or impact, we must also consider the politics of citation and the ways in which

academia reproduces itself through citational practice (Ahmed, 2013; Ahmed, 2017; Burgess et al., 2021). Sara Ahmed (2013) explains how citations are used to create spaces and structures that ultimately form disciplines. These structures are then reinforced by sexist and racist citational practices that reproduce ideas and histories that are centred around certain types of bodies and voices, typically those of white men (Ahmed, 2013, 2017). Wāhine Māori scholars Burgess, Cormack & Reid (2021) also discuss how we, as researchers, are taught to “reproduce settler colonial ideologies” (p. 59) during our academic training, as we are often expected to cite and engage with particular scholarship that has been practically canonised in many of our disciplines through the problematic processes described earlier by Ahmed (2013). If citational practice privileges whiteness so will citation metrics. With this context in mind, we must think about the limitations of bibliometrics in fairly representing or measuring research ‘excellence’, given the realities of inequity in academia, as well as consider the implications of these types of (comparative) measures for female scholars, Indigenous scholars, and other scholars of colour (Locke & Bensky, 2022).

Within all of this, there is widespread documented scepticism of traditional metrics among researchers. A study conducted with researchers and academics at the University of Waikato found “no respondents considered traditional metrics were extremely accurate in reflecting the value of scholarly work”, and a mere 10% selected “quite accurate” (Ferrier-Watson, 2019, p. ii). This response supports an earlier study showing low confidence in the accuracy of traditional metrics in showing the value or importance of researchers’ output (DeSanto & Nichols, 2017). Altmetrics in the same studies fared worse in the views of researchers. It has also been noted that studies like these which examine the views of researchers are relatively rare in the discussion about research excellence measurement (Kondakci et al., 2021).

It has been suggested that the lack of faith in traditional metrics could be attributed to lack of transparency of the metrics themselves (Curry et al., 2022; Yang et al., 2021). Alongside the transparency of traditional metrics is the question of how much researchers understand the metrics that are supposed to represent them. Ferrier-Watson (2019) found that most of the academics surveyed evaluated their self-knowledge of traditional metrics below “not very well”, though academics from the sciences did evaluate themselves higher (p. 51).

For information on metrics, most academics either did not seek out more information or went to Google Scholar. Some of Sarpong's (2021) academic participants expressed how performance measurement can positively hold academics accountable to the taxpayer, and how rankings and metrics can positively influence students and researchers to choose one university over another. The positive and negative themes of their findings were linked to the market-like behaviours of universities, with several academics agreeing that some disciplines are underfunded and undervalued because of this (Sarpong, 2021, p. 121-124).

RESEARCH DESIGN

Research questions

The research project was guided by two main questions:

1. What measures are used to quantitatively measure research excellence?
2. To what extent can these measures best represent research in different disciplines?

However, the data that has been collected from participants is much richer than these questions. As such the decision was made to release this report to inform university practice and researcher support. A forthcoming publication will focus more closely on the research questions initially posed.

Approach

This project, by design, was intended to be exploratory in nature, and used a case study approach to examine researchers' views about a range of quantitative measures of their research impact. The project team carried out a systematic analysis of each researcher's quantitative research metrics across a range of platforms, including traditional research metrics from Scopus, Web of Science and Google Scholar, and also a range of alternative metrics such as usage statistics, policy citations, and social media engagement.

Some metrics participants were asked about in the interviews include:

Metric	Platform	Definition
h-index	Google Scholar, Scopus, Web of Science	"an academic with an index of h has published h papers each of which has been cited in other papers at least h times" (Harzing & Alakangas, 2016, p. 792)
Field-Weighted Citation Impact (FWCI)	SciVal (based on Scopus data)	The ratio of citations received and the average number of citations received by outputs of the same age, document type and field (Colledge, 2017).

The semi-structured interviews were conducted face to face where possible, and over Zoom when meeting in-person was not feasible. These conversations were largely centred around participants' thoughts and opinions about the research metrics we had collected, as well as their wider research impact goals and interpretation of research excellence and research evaluation.

Participants

The participants were all research-active academic staff from the University of Waikato with an established publication record. Participants ranged from early career through to professor. In order to

cover a range of disciplines, including those not generally served well by traditional metrics, one participant was identified from each of the thirteen PBRF panels: Biological Sciences; Business and Economics; Creative and Performing Arts; Education; Engineering, Technology and Architecture; Health; Medicine and Public Health; Humanities and Law; Mātauranga Māori; Mathematical and Information Sciences and Technology; Pacific Research; Physical Sciences; and Social Sciences and Other Cultural/Social Studies.

Data analysis

Interviews were all audio-recorded and transcribed initially using Otter AI transcription software, and corrected by the project team. The interviews were analysed by two members of the project team using thematic analysis. This analysis was then further refined into the common themes outlined below.

Ethical considerations

Ethical considerations were taken into account throughout the study and care was taken at all stages to preserve the anonymity of participants so that they were comfortable to speak freely. In an effort to preserve anonymity the majority of quotes in this report are identified by interview number rather than discipline, with the exception being under Theme 5 where we have identified quotes as coming from the interviewees for the Mātauranga Māori (MM) and Pacific Research (PR) panels. This was done in order to give context to discussions around cultural considerations and cultural labour in academic environments. Information has been redacted from direct quotes where it may have identified the speaker.

A research data plan was created and followed with additional efforts made to comply with Māori data sovereignty best practices, including ensuring data was stored in secure local servers and transitory data was deleted promptly. This research project was approved by the University of Waikato Human Research Ethics Committee (HREC2022#01).

Limitations

One researcher cannot represent an entire field, and as such it is important to acknowledge that the research is exploratory only and results cannot be generalised. While selecting one participant from each panel allowed us to ensure a spread across disciplines, the spread of panels across the disciplines does not align with the number of staff employed at the University of Waikato, which again enforces the need to avoid generalisation of these views.

THEMES

Theme 1: No one-size fits all: Impact and excellence varies greatly across disciplines and matter more to some disciplines than others

Diversity of the scholarly landscape

Researchers were quick to point out the differences in how research evaluation is carried out between disciplines:

“some people will get represented in Scopus better than others ... there's different ways. There's not a one size fits all.” (Interview 7)

Because there's no such thing as the perfect metric, the alternative is to harness various ways of measuring different aspects of research, essentially stitching together the figures to provide a “snapshot” of a researcher's output:

“Because we publish in different places, and seek different types of impacts, I think we need different measures.” (Interview 7)

However, it would be impossible to find a single point of similarity with which to judge them. For example, it would be unfair to compare a career researcher who has been consistently publishing highly-cited works in high-status journals for fifty years with a newly fledged researcher in a field which does not have published outputs (e.g. creative performances). The only thing these two hypothetical persons may have in common could be their place of work. In all other measures, they are at opposite ends of the spectrum or on a different spectrum altogether:

“The mathematicians would weight journal publications very highly, because their journal publications have a lower publication expectancy rate, so the mathematicians would say, if you publish two good journal papers a year, you're doing really well. In computing, some will say, well hang on, you need to be publishing at least four papers a year, but they're probably conference papers. ... So it's different and that's often been a bit of a challenge.” (Interview 9)

“I think that there's a lot of ways to characterise, scholars are different even within well across the panels, certainly, but within even our [redacted] panel, we're different types of people.” (Interview 7)

“I've had a slow start. I had a period ...[of] almost non existent research years as chair of department. ... And all of my growth in my h-index has been since that time.” (Interview 6)

Wide range of metrics in use

Currently, a range of metrics are employed in an attempt to provide different ways to look at researchers. Publication count is one, with a higher output generally perceived as better - with some exceptions (Ansele, 2023) or subsidiary concerns for published papers such as the quality of the

journal in which research is published, the speed from discovery to publication, the number of times research was viewed, downloaded or cited, and the quality of those views, downloads or citations. For some researchers, their work is highly targeted to other academics; for others, reaching teachers or on-the-ground practitioners of research may be more important; for others, reaching a community group which might benefit from the research shapes their emphasis on these metrics. Impact can also be measured through influence on government policy, or through mentions in media outlets such as TV, radio, newspapers or online:

“You know, as engineers [we] want to bring positive changes to society. Has this been achieved? And again ... this does not depend on how many patents you have, it does not depend on how many papers are published, how many citations you attract. All these are actually useless.” (Interview 2)

“At my stage of the career, I don’t have to publish huge numbers of things, I’m more interested in doing quality and things that are policy relevant ... and relevant on the ground to people.” (Interview 6)

Knowing the limitations in comparing across disciplines, there are some attempts to counterbalance this with metrics like the FWCI creating a more fair comparison. However, this only works if research falls neatly into Scopus’ predefined categories, which many research outputs and/or researchers do not:

“... we're always folded into something else instead of being a category in our own capacity, you know, so if they had citation indexes that recognised Indigenous scholarship as its own thing.” (Interview 12)

Limitations and implications of using metrics to assess researchers

The metrics on offer do not take into account such differences as career stage progression, life experience, age, ethnicity, gender and other factors which may influence how well a researcher does when evaluated. The result is a “standard” which may derive from a historical evaluation of a field of research, which naturally will unfairly weight some of these factors:

“I think if there was a multiplier effect that could be added to a citation index in recognition that this was a Māori, Pacific and Indigenous piece of work, and so it was citation times three, because of the extra effort it takes us to get the interest of a journal, to get through the reviewer feedback system, which is strongly weighted to the Western cultural paradigms and mainstream themes.” (Interview 12)

We find that despite many metrics being used, most researchers consider them limited and only mildly of interest. Researchers are also strongly opposed to their worth being measured using metrics, particularly without their input. The push for qualitative metrics at least includes the researcher in the conversation and allows them to contextualise why they may have done well on some metrics and less well on others. But researchers already have many demands on their time and little inclination to learn about the metrics which they do not feel represent them well. Many of the comments received in the interviews indicate that researchers evaluate their work based on their instinct rather than by a metric:

“I don’t think these metrics are that useful, because you know whether you’re happy with your work and are satisfied with your work.” (Interview 4)

“Anyone who performs knows in the moment of performing what impact they're having on people.” (Interview 1)

Researchers seemed resigned that abandoning metrics altogether would be difficult in the current research environment:

“I think metrics are always going to be used, because it is a simple way for people to evaluate themselves against other people and a simple way for panel members to evaluate themselves. ... I just think it’s really important to be aware of limitations, and it just doesn’t work in some areas.” (Interview 3)

“It would really [be] nice just to have a standardised metric, so that we can become familiar with it, then glance over and understand everything that we've seen. But the problem is how to do that. How do you sort of equalise the different fields and everything like that?” (Interview 8)

Theme 2: Lack of understanding of metrics and reliance on H-index

There was a general feeling among the interviewees that not only were metrics problematic in their application, but that they were not well understood. Researchers tended to find a metric that they liked or recognised and stuck with it (even though, in some cases, another metric might give them a higher score).

Lack of understanding around metrics

Across all the interviews there was a lack of prior knowledge about many of the metrics presented in the Research Profiles Health Check. Even after being provided with explanations many interviewees expressed that they did not understand how some of the metrics were calculated, or the significance of particular numbers.

Multiple participants mentioned that even if they used certain metrics in their PBRF, those assessing the portfolios would not understand the significance of many metrics. The idea was expressed that the PBRF panels should not evaluate researchers based on metrics without having a thorough understanding of them and their limitations, for example:

“Would that mean anything to the panel?” (Interview 1)

“In terms of metrics on a panel, what you'd have to do is make sure the panel members totally understand how the metric works, understand what the failing of the metric is, and also has a way to make it relative.” (Interview 10)

More likely to use h-index because they understand it

The h-index was the only measure that was widely understood and many participants expressed the opinion that they used the h-index because it was a simple, transparent calculation. That said, some

participants admitted they either did not understand the h-index, or did not believe it was relevant or useful:

“H-index is not that useful. ... You could say this applies to a lot of these metrics.” (Interview 1)

“I've heard of the h-index. I know that, in conversations in other disciplines, people get very animated about it. I'm less animated, because I don't fully understand how it works.” (Interview 13)

A number of participants expressed awareness of various limitations of h-index as a metric, however despite these limitations, many participants mentioned they would be likely to still use h-index over other metrics:

Interviewer: If you had to choose between using a field weighted citation impact or your h-index, what would you use?

Interviewee 6: “Oh, I mean, the h-index because I, that's what I understand. But I guess if I now reflect on that field weighted index, I guess, I need to think about what that tells me in terms of writing my PBRF folio. Trying to weave that in and tell the story around it. So that will take a bit more thinking.”

Lack of knowledge and trust around field weighted citation metrics

The interviewers asked participants for their opinions and feelings about a range of metrics from Sci-Val, in particular the FWCI. There was very little prior knowledge or understanding of this particular metric, but also the idea of field normalised metrics seemed to be new to many participants. Once the metric was explained to participants some felt it was still not relevant, especially the participants from the Indigenous, arts, and humanities subject areas:

“I'm not expecting Scopus to be my place.” (Interview 5)

“It's like a different language, it really is.” (Interview 11)

“I'm not sure how I would use it.” (Interview 1)

After the FWCI was explained to them, a number of participants from the STEM subjects reflected that they could see the benefits to field normalised metrics, such as FWCI. Despite these admissions, there still seemed to be a mistrust around the complicated calculation of this type of metric. There were a number of concerns around how a field was determined:

“I guess my biggest concern about those metrics, about that type of metric is having to define the field because by its very nature, the work I do is cross-disciplinary and multidisciplinary.” (Interview 3)

“So I guess if it's field normalised, right? Probably is more accurate compared to a h-index. Again, it is very hard for me to give accurate statement, whether that is good or not, because I know it's such a narrow field is about how they come up with this number.” (Interview 2)

“I know how to interpret an H index and what is a good one versus a bad one. ... Whereas I had no idea how to interpret the field weighted. ... I don't know the scale to understand how to interpret it.” (Interview 7)

It is interesting to note that while a number of participants were aware of shortcomings of the h-index, only one participant mentioned the need to use a range of metrics or the need to combine quantitative and qualitative approaches to assessing research excellence, as has been recommended in the responsible metrics movement (Curry et al., 2022; Hicks et al., 2015).

Views on differences between platforms

In multiple interviews there was confusion around the differences between metrics available across different platforms, such as Web of Science, Scopus and Google Scholar. Multiple participants talked about whether one platform might be more accurate than another, or expressed a preference for Google Scholar as it produces higher citation metrics and better coverage in some disciplines:

“Is Scopus more accurate, or is Google Scholar more accurate? I just like the numbers better in Google Scholar.” (Interview 7)

“I probably would prefer Google Scholar over Scopus, because it is more all encompassing, and I think it includes book chapters. Am I right in thinking that? Whereas it's more widely inclusive.” (Interview 3)

Theme 3: Metrics are subjective and can be gamed

General scepticism and caution around the use of metrics

While many researchers claimed to understand the need for metrics and measuring research impact, there was a lot of concern about how subjective they can be. One interviewee expressed a lot of scepticism about the use of metrics indicating:

“once you tear away all the framing around it, metrics are subjectively defined, manipulated, calculated.” (Interview 1)

The interviewee further maintained that “real impact”, is measured outside of the academic sphere with social justice and environmental effects. This view on the subjectivity of metrics was echoed in Interview 10 when discussing metrics and how publishing less often in higher ranked journals produces a better rank than publishing more often in lower ranked journals.

Another interviewee considered the measuring of their research to be viewed simply as “other people’s perspective”. When discussing specific publications that get more attention, they had “no idea why one would get picked up over another, you know, in terms of interest” suggesting that the popularity of a publication was not always indicative of its importance or value to the author. The perceived subjectivity of metrics is linked to gaming the system. They elaborated by saying that:

“I'm not sitting there trying to strategise how this is going to push my citation index up or my h-index up. But maybe some people do.” (Interview 12)

More prominent metrics are often associated with certain databases which created tension for some interviewees:

"I'm not always sure where I fit in Scopus. I'm not really science-y. I kept hearing: "yes, you should, you need to be in Scopus." Okay, I'm in Scopus, but I'm not sure how I fit there as ranked against other people. And whether they're looking for someone like me" (Interview 5)

Gaming of metrics and impact on career

Some interviewees felt that because promotions were granted based on metrics, there was a temptation to "play the game", despite not wanting to detract from the research. They also commented on the positive correlation between co-authors and metrics. It was suggested by some that co-authorship did not always reflect personal contribution to a publication.

"I think it's becoming easier to get higher metrics, as people tend to collaborate more as you get papers with larger numbers of co-authors on them." (Interview 4)

Metrics measure different things at different times, and a researcher's focus changes over time:

"there's different ways of breaking these metrics up. I'm not saying that you need to be the lead author on every paper, that's maybe my focus at the moment, and that will shift over time." (Interview 4)

Other interviewees expressed an internal conflict between recognising the propensity to game metrics and feeling pressured to participate in gaming behaviour:

"I don't particularly want to go down the cynical path of gaming the system, but I do want the work that I do to have some recognition...It's a game. And so a lot of this is about a system that can be gamed." (Interview 11)

"It's a numbers game. Whenever we are using this as a driving force for performance evaluation, people will try to cheat." (Interview 2)

Citation counts and h-indexes do not take into account contribution toward a publication; they don't show whether an output is sole-authored, or if an author is 50 of 100 contributors. When participants were asked about how they felt about the metrics of their top cited papers, overwhelmingly the outputs they were most proud of were those they had led or that they felt had made an impact beyond citation count:

"I don't think of it as my own publication. But obviously the way the metrics and citations work, that doesn't matter." (Interview 6)

"I'm kind of very proud of that paper, potentially more so than some of the other ones on that list. Because it had impact, a real impact, like it's actually being used for decision making. So yes, so it's that balance of, you know, how do I get my numbers up? How do I get high-impact papers versus what's useful?" (Interview 4)

Theme 4: A general nervousness around metrics or disdain, especially for Altmetrics

There was an overriding sense of anxiety about metrics, with one interviewee explicitly stating "they make me nervous" (Interview 5). While interviewees did not have an in-depth understanding of the

limitations of various metrics, they were all vaguely aware of the potential for manipulation, or “gaming”. In a few cases the responsibility for understanding and monitoring this was placed with university leadership:

“I feel that as individual researchers, they shouldn’t really care that much about the metrics, and also the managers shouldn’t use these numbers to judge or evaluate individual researchers’ performance.” (Interview 2)

“To me, it means that people who are in charge of, are in leadership positions, managerial positions for making decisions on promotion, need to be really aware of the limitations of these metrics.” (Interview 3)

Interestingly, the strongest scepticism was around altmetrics, and social media metrics in particular. There was a perception that if researchers did not use social media, then this was irrelevant for them. There was little to no acknowledgement that social media metrics could accumulate without their active participation. 6 out of 13 interviewees specifically stated they did not engage with social media when asked about their thoughts on altmetrics. Despite this, some interviewees felt that either they should engage more with social media “it’s probably something that I should be better in” (Interview 4), or that they were being pushed towards social media “we’re being pushed all the time to engage with social media. ... I don’t really want to go on Twitter and social media. I need a young person to do that” (Interview 3).

There was a sense that altmetrics could be manipulated (more so than citation-based metrics), and were more prone to be influenced by what was, or could be shaped to be news-worthy:

“If you write something really inflammatory, and controversial, and you court that controversy, then you will be cited all over the place widely for years.” (Interview 1)

Overall, interviewees felt that metrics in general were a poor proxy for quality, and generally lacked confidence that they would be used in the “right” ways:

“I also hope that we don’t get too caught up in metrics of ways to evaluate people. That they are one tool, and that there are so many other ways when evaluating even just research quality, let alone the overall characteristic of academics. [If] we’re talking about PBRF panels there are so many ways to evaluate and just not get too caught up in pigeonholing everybody into one, which is the easy thing to do just because it makes their life easier, the evaluators.” (Interview 7)

Theme 5: Cultural considerations and the importance of the collective

Compromise and publishing practices

A recurring point of discussion throughout the Mātauranga Māori (MM) and Pacific Research (PR) panel interviews was the compromise that takes place when producing work or research that satisfies the assessment criteria of the PBRF panels. For example, the MM interviewee felt that the push to submit research to ‘highly ranked’ global journals (that usually carry a heavier weighting in research assessment) often means that Indigenous scholars essentially have to dilute their work, to either make it more palatable to a wider audience or to satisfy the wants of the reviewers of these

journals who themselves may not have the cultural expertise to understand or appreciate the nuance and context of Māori-focused research:

“It's disheartening the kind of feedback we get from reviewers who show very little understanding, they're just more concerned with how is this going to appeal to a CEO in Europe or something like that...” (MM)

“Most of our work has to get oriented as much as possible to a wider audience, so that comes through the reviewer feedback. So when working at that level of journal, a lot of iterations to make sure that it's got this wider acceptability or readability for a global audience. Now, that's not the case with MAI or AlterNative, we can stick within our knowledge paradigms.” (MM)

The PR interviewee expressed similar tensions and further explained that ultimately, they write for their own community of researchers and practitioners whose feedback is vital in their own research processes:

“...the pressure that all Pacific researchers are under, trying to measure up to getting that peer reviewed international journal article. It's just so, kind of drilled into you from the time that you publish your PhD and on, that you could feel kind of like a failure because your CV doesn't reflect all the hard work that went into a three-year research project like that, versus a journal article that garners a lot of citations ...” (PR)

“I do want it to get back to the end users, who are Pacific people, and teachers and [redacted]. And I want those groups to see that we're represented in the research. I want them to be academically rigorous. And if it's absolute rubbish, and I don't know what I'm talking about, they would be the first ones to tell me that .” (PR)

Conflicting objectives and motivations for research

For the MM interviewee, there is significant tension between the desire to write in order to serve the needs of their scholarly communities and writing to appease journal requirements and reviewers. They also noted that they have more flexibility to move within their own knowledge paradigms when submitting their work to local journals that have experience with presenting and publishing work that centres an Indigenous worldview, however those journals are often not as highly ranked or heavily weighted when it comes to researcher assessment. Further to the PR interviewee's comment on garnering citations, the MM interviewee notes that the use of metrics as an evaluative tool for research excellence can be limiting and incompatible with their objectives as a scholar, stating:

“We're not writing for scope and scale, we're writing for quality of a knowledge system, we're writing because we want to bring that [Indigenous] voice forward. Now, if I was going to be really tactical I would abandon all of that and just write for scale and write to have the most broadest citation marketplace appeal—so as many consumers of my work would be interested in citing it. So, to me, that's a fundamental tension in the way that citations are used to grade a scholar's quality of work.” (MM)

Making contributions to their field of knowledge is a fundamental objective for an academic. The emphasis on quantitative measures, such as citation metrics, in the assessment of researcher performance can often undermine that scholarly objective when researchers feel they have to forgo producing work that will feed directly into their communities in order to produce work that will contribute to their own career progression. As demonstrated by the MM interviewee, this can be a

significant conflict for Indigenous academics who are often community driven and focused. These concerns are also discussed by the PR interviewee:

“Without trying to represent all Pacific researchers, I feel like the pressure on Pacific researchers is that you, at any given moment, do feel like you have to represent your culture. Those cultures teach us to make the sacrifice, and I don’t want us to change that. That’s a wonderful part of our culture. But what it does is then put you in tension with a system that is about being very selfish, and individually focused on just your career ...” (PR)

Cultural labour not captured by research assessment measures

In addition to deriving meaning from being in conversation with other Indigenous scholars, the MM panel interviewee also notes the reality of the amount of cultural labour that goes into creating space for Indigenous voices in their work, especially when they are situated in a discipline that has not traditionally centred Indigenous perspectives or lives:

“Sometimes I look at the citations and see who's cited it and I feel good when I see an Indigenous scholar has cited it. Because I feel like it's hopefully affirming and reinforcing what they're doing, and that we're all helping each other, in a sense, to advance this body of work collectively.” (MM)

“And so, a book chapter, a book itself, a special issue, all of those things come from a sense of collective and community orientation for the community of scholars, not for one's own self promotional angle or, you know, advancement, that generally speaking I'd like to see those recognised and given weighting as well—say, maybe someone in the arts will get it for a performance—we should get it for the mission that's involved, not just for being self-serving, and for the fact that it's amplifying impact in community.” (MM)

These also feed into comments around mentorship and the foresight of Indigenous academics in ensuring there is adequate support for junior scholars coming up from their postgraduate studies into the academic workforce, and also how that substantial labour is not fairly represented in their assessment as a researcher. The MM interviewee notes that for Indigenous scholars, “it’s a small pipeline” of students gaining PhDs in their subject area and there is a lot of cultural as well as administrative labour that goes into providing opportunities for these up and coming researchers:

“...I might have my PhD students, invite them to write book chapters with me or solo, that's not recognised in the PBRF system. But for an up and coming field of Indigenous scholarship and Indigenous scholars, it's a really important outlet, but there's no kind of—because it's not peer reviewed, and so the rigour's considered less than being in a dean's list article or whatever journal. So, I've always felt that, when I try and portray my narrative and explain that, it falls on deaf ears.” (MM)

“And also mentoring relationships to get—like, I know at the moment there's a lot of work taking place to get as many students again to the [redacted] and then into an Indigenous program of mentoring where senior academics will mentor younger academics and on a PBRF form that just will be a line or two, but the amount of work that's going into supporting the development of these kind of precious little taonga that are spread all over the world and help them get ahead and to feel like they've got a voice - I think it's not captured...” (MM)

The PR interviewee also described how this commitment to nurturing and supporting Pacific students in the academy ultimately influenced their choice of career pathway:

“...my choice was to take my PhD and run with it and turn it into a book or step in and be [redacted] at a time when our Pacific students would have suffered if somebody didn't step into that role.” (PR)

Theme 6: Disconnect between what they value and what they think their peers value

There was a general tendency for interviewees to describe their goals, aspirations and strategies as separate from others'. Some of these differences relate to discipline, some to career stage, and some to the intended impact or audience of an output.

Interviewees commented extensively on what colleagues valued or practised, with the implication being that the interviewee themselves did not value the same things. Self-promotion, for example, was often posited as something “other people did”. One interviewee commented that “everybody wants to be kind of like a big shot in Google Scholar, like they wanted to become a big shot in Tik-Tok” (Interview 2). Another stated:

“There are people that are very good at that media outreach. There's nothing wrong with that. I think that's fine. ... It's a lot of time, a lot of effort to do that properly, to do that well, and I haven't decided to put my efforts into that.” (Interview 4)

Occasionally these efforts were referred to in vaguely negative terms, one interviewee described this sort of self-promotion as “humble bragging”. In other cases, interviewees felt that others were putting in more effort, and that they themselves could be handling their work more strategically:

“It's not something that I pursued strongly. I know some of my colleagues do sort of work hard at making sure their publications are, you know, appropriately accessed. And it's not something that I've done. It's been on my to-do list for several years, and I've not got around to doing...” (Interview 9)

Interviewees' perception of themselves often stood in opposition to using or valuing metrics, which they said they did because the University required it, either for rankings or promotion, rather than being something they valued themselves. There was clearly a tension between what they wanted to achieve as a researcher and how they wanted to achieve it, and what they felt they needed to do in order to be employed in academia. One interviewee, when asked about metrics, responded, “They're important to the university. They're not important to me.” (Interview 3).

Theme 7: Publishing decisions are driven by a diverse range of factors

The outlets or places in which researchers choose to publish their work can directly influence their research impact.

Audiences come first

A recurring point made by interviewees was that audience was their first consideration when making publication choices. Researchers think carefully about where they publish their work and in what form (e.g. conference paper, article or book) and these decisions are guided by the intended audience. Interview 11 mentioned, "I do think about audience with where to publish, like, who needs this, who is going to read it?" Interview 3 also said when it came to disseminating their work, they usually tried to publish it "in the place where it's most likely to be used by people." Interview 8 pointed out:

"I have always sort of tried to think about who reads these journals, who is going to benefit the most by publishing here. So without thinking about the metrics, and the metrics is more of an afterthought." (Interview 8)

Relationship with publishers

Throughout the interviews it was found that relationships with publishers oftentimes influenced where researchers chose to place their work. For example, Interview 1 told us that they consider where publishers are based (geographically) and how that might affect the reach of any work published with them. Indigenous researchers felt an obligation to publish in certain spaces and felt that their relationships with local journals are more reciprocal. The journals get to publish local Indigenous scholarly work, but in return, the researchers are provided with a culturally competent space where they do not necessarily have to "hold back" or try and tailor their writing to a global audience who do not understand the specificities of our context here in Aotearoa New Zealand. Interview 12 discussed this dynamic:

"So those relationships are also important. And knowing that it's likely to get through to reviewers who have got a cultural responsiveness, even if they're not Indigenous, that the feedback we get is respectful, honouring and understanding versus some of these other journals." (Interview 12)

Whereas Interview 3 mentioned the relationship between publishers made them feel like there was an "in club of people" or "a closed shop". If you are not in that club, it is very hard to get published in certain highly esteemed journals.

Journal impact and reputation

The quality or status of a journal is another consideration that drives researchers' publication decisions. While many interviewees said they "just know" which journals are best for them to publish in, they still made mention of taking into consideration the journal rank and reputation. There was also a general understanding that publishing in highly ranked journals would potentially improve their research impact and reflect favourably on the way they are assessed. Interview 6 stated, "We were aiming for quality international journals ... for me, out of that is that we are not publishing for publishing's sake, we are publishing significant results." Interview 10 expressed that they often went for journals that had the biggest impact because metrics were "good for [the researcher's] ego".

Some interviewees also mentioned the importance of publishing for their students, though acknowledged there was not as much pressure for them to get into the top-tier journals. Speed of publication

was also critical, as it can sometimes take years to get published in an international source, and as such some interviewees commented that a local source may be preferred in order to get data out there before it is “stale”.

Theme 8: Varying levels of importance in getting research out to the public including OA

Incongruence of values and behaviours in the context of open access

Overall, researchers expressed varying levels of understanding in relation to communicating their research to the public as well as making their work publicly accessible through open access initiatives. The degree to which this was important to them also varied among interviewees.

While the majority of interviewees agreed in principle that open-access publishing was important and expressed support for it, less than half of the participants had made a deliberate effort to primarily publish open in their own careers. Researchers who strongly valued open access had research goals that included influencing public decision-making and social policy:

“What my research goals are ... I guess the aim is to continue to provide science that's useful for society.” (Interview 4)

“There's several primary goals. So if you ask me, what am I trying to accomplish my research, it's to have research impact. Impact on communities and societies.” (Interview 7)

“So I like to do research that makes a difference, but also do it in an academically rigorous way.” (Interview 3)

Researchers mentioned an array of pathways to communicate their research to the public, either directly through open-access publications, media/social media outreach, conferences and other communications like seminars or productions. There was a tendency to support the open attributes of research profile tools and metrics. Notably, the potential that attributes like discoverability, access, and impact had towards helping them achieve their own goals.

Interviewees said that ideally, their research publications were published open access and in journals that they trusted and were familiar with, but financially this was not always a viable option. Some interviewees prioritised the quality of journals they published in regardless of opportunities for open access. Some also accepted open access as an expectation of them as part of a publicly funded institution. For interviewees that published open access, the primary method to pay for open access fees was by using research trust accounts and grants when available:

“So, yeah, I definitely think it shows that that you can get more readership with open access, and there's an ethical aspect to that.” (Interview 11)

“I support open access, I feel that it's a good way to publicise your work.” (Interview 2)

“If you've got a research grant, then the research grant can pay for it. If you haven't got a research grant, that becomes a little more challenging.” (Interview 9)

“I would like it open access, but that cost two and a half thousand Euro ... I guess if I had thought about it at the time, ... I would have put money aside.” (Interview 13)

Awareness of the ability to deposit their work directly into open repositories such as the university's Research Commons varied among interviewees but was positively recognised throughout as a beneficial access point to their research:

“...I'm excited by that, because that makes the work more accessible.” (Interview 6)

“The idea that we have it available on Research Commons so that it's accessible by people can download it, I think, is a really important thing.” (Interview 9)

There were some barriers that held interviewees back from depositing their outputs into Research Commons including reservations about copyright, time pressures, and confusion around shared authorship. Those that viewed the depositing process as achievable incorporated it as a part of their communication and outreach. Negative experiences with the process coincided with researchers that had not yet assessed the eligibility to deposit multiple publications and were unsure of copyright limitations with publishers and coauthors:

“Sometimes these are other people's articles that we're trying to share. And I don't have control whether they've stuck it in the commons or whether they did open access, so I don't.” (Interview 7)

“It seems somewhat complicated, given the amount of stuff that I had put out already, to work through all of that.” (Interview 13)

Communicating and disseminating research

Interviewees made a clear distinction between communication within academia among research peers, and outside of academia with practitioners and the public, often referring to speaking different languages and reformatting their work for different audiences:

“It's also about putting it out in the public arena for people to try and understand, for the general public to try and understand what people like I do and maybe also consider some of the issues that I think about.” (Interview 10)

“...not necessarily through Twitter, but condensing quite dry, not dry. I mean, my research isn't dry, but difficult to follow, if you're not an expert in the field, making it accessible to people, I think that's a good thing to be doing. And that does generate impact.” (Interview 4)

When analysing participants' methods of communicating their research, we found that they were the most divided when asked about metrics that measured social media and media activity. Some participants felt that social media and media communications were powerful mediums to convey research to the public. The definition of 'public engagement' meant different things among interviewees and for some, it meant achieving research goals:

“I think it's a really good way of looking at maybe general public engagement with research, and it's possibly something that as researchers we all need to be aware of.” (Interview 3)

“My personal view is that this sort of information could be potentially more useful than h-index. ... Because sometimes the way to measure the impact of your work is really for example, has it attracted some kind of media coverage? Or has it been commercialised? How has it been applied in actual practice?” (Interview 2)

“I'm excited by the number of media opportunities I got last year. I mean, they don't appear in metrics. But to me, that's a sign of visibility and relevance.” (Interview 6)

Social media as a means of dissemination

All participants but one shared their lack of engagement with social media, stressing that a lack of knowledge, confidence, time, and training were barriers to utilising the platforms as mediums of research communication:

“Some training on how to do it would be useful. Albeit, it's just there's a limited number of hours in the day. ... You have to decide which bits of this research process do you engage with?” (Interview 3)

“I don't tweet and I don't Facebook really. ... I can kind of see how you can widen your audience that way, which is, I think, a good thing. But there are people that are very good at that media outreach. ... They get a lot of traction with their work. It's a lot of time, a lot of effort to do that properly, to do that well, and I haven't decided to put my efforts into that. So I don't really focus on those metrics. Because I haven't put time into them.” (Interview 4)

“I have never tweeted, I don't have a Twitter account. No TikTok, no Instagram, I have a Facebook and I've never posted a thing on it.” (Interview 7)

Some interviewees did not see the value of social media when it came to disseminating their research. There were some reservations about social media including comments on safety concerns, the potential dangers of excessive social media use, and implications of algorithms. It was suggested that social media metrics had the propensity to be gamed according to multiple interviewees, particularly when large research groups and affiliations are involved:

“I feel that researchers shouldn't be like an influencer, you know, in Tik-Tok or YouTube, you know, you want to see how many followers you have, how many clicks. If you look too much [at] that you can't do a good job.” (Interview 2)

“I'm very nervous about how to use that successfully to avoid things like pile-ons from right-wing extremists, things like that. I once found my name on an alt-right website.” (Interview 5)

Overall, the support for aspects of open access suggested that interviewees highly valued the connections their research made with audiences, and that openness helped create these opportunities. Interviewees equally valued their autonomy to do their work and to share it on their terms. However,

there were some barriers to achieving this balance and this was largely due to participants not knowing or understanding how to utilise tools available to them, with many also being disinterested and others cautious.

CONCLUSION

One of the main sentiments that came to light in the interviews was that our researchers care deeply about what research they are doing and why they are doing it. Researchers easily articulated their research goals and intended impact, none of which had anything to do with personal gain. But it is clear that they are operating in an environment which creates tension between the good they want to do in the world (arguably, the reason for the existence of universities), and the things they need to do in order to sustain an academic career.

The implication is that metrics and the way they are applied ultimately act as a barrier to impactful research and are a source of stress for researchers. Responses made it clear that for Indigenous and other marginalised scholars, the systems as they stand actively erode traditional knowledge systems and instead uphold norms grounded in Western (and often male) worldviews. In the context of Aotearoa New Zealand, it might be said that our current research culture is antithetical to the principles of Te Tiriti.

This paper does not argue for an end to research evaluation or that there is no place for metrics in the scholarly ecosystem. Rather, that metrics should be considered one tool amongst many which researchers might choose to use to tell the story of their research and its impact, and those who evaluate research should value them accordingly. In order to facilitate this, there needs to be greater awareness more generally of what metrics are on offer, how they work alongside non-quantitative measures, and an understanding of their limitations. This will align us with the Ministry of Business Innovation and Employment (MBIE) interest in “approaches to assessment of research that [rely] less heavily on bibliometrics” (MBIE, 2022, p. 11). We suggest that the best way forward is to evaluate the metrics; that is, list those metrics that we measure with a note of whom they best serve and who is disadvantaged. We should be searching for a way to incorporate existing metrics where appropriate, while acknowledging their limitations. Simultaneously, the range of metrics and indicators available should evolve over time, allowing newer indicators to be implemented and less helpful ones to be actively discouraged if they are found to be wanting.

Ultimately this also becomes a question of how metrics are integrated into the systems and processes that drive academia, such as PBRF and other research assessment exercises, promotions and rankings. It is these features of the scholarly ecosystem that incentivise certain behaviours. The Metric Tide (Wilsdon et al., 2015) refers to this as “cultures of counting” (p. 79). This sentiment is echoed in the Leiden Manifesto “The problem is that evaluation is now led by the data rather than by judgement ... we risk damaging the system with the very tools designed to improve it” (Hicks et al., 2015, p. 429). In addressing the systemic issues in academia, it is worth thinking about who benefits from the system as it stands. For example, large publishers profit from a scholarly ecosystem which values prestige over access and quantity over quality.

This change can be viewed as part of a wider development of research culture which “encompass[es] the behaviours, values, expectations, attitudes and norms of our research communities. It influences researchers’ career paths and determines the way that research is conducted and communicated” (Royal Society London, 2023). Systemic change is not easy, but steps can be taken by individual institutions on a number of levels. For example, becoming a signatory to the [Knowledge Equity Network](#) or the [San Francisco Declaration on Research Assessment](#) (DORA), or a commitment to abiding by the [Leiden Manifesto](#) principles. These on their own will not necessarily trigger change, but can become the catalyst for a move towards a research culture which is better aligned to our institutional mission. The University of Melbourne, for example, was the first Australasian university to sign up to DORA in 2020. Queensland University of Technology followed in their early adoption of the [INORMS More Than Our Rank](#) initiative. In a 2021 survey on responsible metrics use, 40% of responding institutions had either signed DORA or were likely to sign, up from 30% in 2019 (Price, 2022). Similarly, the University of Bath established their own principles of research assessment and management in 2017, among others (University of Bath, 2017). Of course, signing DORA or having a responsible metrics use statement on its own will not fix what is broken in the system. But such steps can act as a catalyst for a change in culture. We know from our interviews that in order for researchers to change their behaviours in good faith, they need to know that other factors in the system (bibliometric practitioners, colleagues and leaders) are acting according to similar principles as themselves. For this reason, change needs to start with wide-ranging conversations and visible buy-in.

The upshot of this research is that metrics made good servants but bad masters—when we value what we can measure rather than measuring what it is we value (HuMetricsHSS, n.d.), we incentivise behaviours that optimise the numbers, rather than the outcome. With perverse incentives removed, we can allow our researchers to realign their research practices with the values that led them to academia and to work within a university. When they are recognised and acknowledged as individuals within their different disciplines and cultural settings, they can be allowed to tell the story of their work in ways that illuminate their unique value.

REFERENCES

- Ali, M. J. (2021). Questioning the impact of the Impact Factor: A brief review and future directions. *Seminars in Ophthalmology*, 37(1), 90-95. <https://doi.org/10.1080/08820538.2021.1922713>
- Anae, M. (2016). Teu le vā: Samoan relational ethics. *Knowledge Cultures*, 4(3), 117-130. <https://www.proquest.com/scholarly-journals/teu-le-va-samoan-relational-ethics/docview/1906046030/se-2>
- Ansedé, M. (2023, April 2). One of the world's most cited scientists, Radael Laque, suspended without pay for 13 years. *El País*. <https://english.elpais.com/science-tech/2023-04-02/one-of-the-worlds-most-cited-scientists-rafael-luque-suspended-without-pay-for-13-years.html>
- Ahmed, S. (2013, September 11). Making feminist points. *Feministkilljoys*. <https://feministkilljoys.com/2013/09/11/making-feminist-points/>
- Ahmed, S. (2017). *Living a feminist life*. Duke University.
- Asaolu, O. S., Jaiyeola, T. G., Usikalu, M. R., Gayawan, E., Atolani, O., & Adeyemi, O. S. (2022). U-index: A new universal metric as unique indicator of researcher's contributions to academic knowledge. *Scientific African*, 16, Article e01231. <https://doi.org/10.1016/j.sciaf.2022.e01231>
- Bakker, C., Cooper, K., Langham-Putrow, A., & McBurney, J. (2020). Qualitative analysis of faculty opinions on and perceptions of research impact metrics. *Colleague & Research Libraries*, 81(6), 896- 912. <https://doi.org/10.5860/crl.81.6.896>
- Bornmann, L. (2017). Measuring impact in research evaluations: A thorough discussion of methods for, effects of and problems with impact measurements. *Higher Education*, 73, 775-787. <https://doi.org/10.1007/s10734-016-9995-x>
- Bornmann, L., Haunschild, R., & Adams, J. (2019). Do altmetrics assess societal impact in a comparable way to case studies? An empirical test of the convergent validity of altmetrics bases on data from the UK research excellence framework (REF). *Journal of Informetrics*, 10(1), 325-340. <https://doi.org/10.1016/j.joi.2019.01.008>
- Buckle, R. A., & Creedy, J. (2019). The 'disciplinary effect' of the performance-based research fund process in New Zealand. *New Zealand Economic Papers*, 54(2), 107-126. <https://doi.org/10.1080/00779954.2019.1636122>
- Burgess, H., Cormack, D., Reid, P. (2021). Calling forth our pasts, citing our futures: An envisioning of a Kaupapa Māori citational practice. *MAI Journal*, 10(1), 57-67. <https://doi.org/10.20507/MAIJournal.2021.10.1.8>
- Carpenter, C. R., Cone, D. C., & Sarli, C. C. (2014). Using publication metrics to highlight academic productivity and research impact. *Academic Emergency Medicine*, 21(10), 1160-1172. <https://doi.org/10.1111/acem.12482>
- Chatterjee, B., Cordery, C. J., De Loo, I., & Letiche, H. (2020). The spectacle of research assessment systems: Insights from New Zealand and the United Kingdom. *Accounting, Auditing & Accountability Journal*, 33(6), 1219-1246. <https://doi.org/10.1108/AAAJ-01-2019-3865>
- Colledge, L. (2017). *Snowball metrics recipe book: Standardised research metrics: By the sector, for the sector* (3rd ed.). Elsevier.
- Cooper, I. D. (2015). Bibliometrics basics. *Journal of the Medical Library Association*, 103(4), 217-218. <https://doi.org/10.3163/1536-5050.103.4.013>

- Csiszar, A. (2017). The catalogue that made metrics, and changed science. *Nature*, 551, 163–165 (2017). <https://doi.org/10.1038/551163a>
- Cupples, J., & Pawson, E. (2012). Giving an account of oneself: The PBRF and the neoliberal university. *New Zealand Geographer*, 68(1), 14-23. <https://doi.org/10.1111/j.1745-7939.2012.01217.x>
- Curry, S., Gadd, E., & Wilsdon, J. (2022). *Harnessing the metric tide: Indicators, infrastructures & priorities for UK responsible research assessment*. Research on Research Institute, The Metric Tide Revisited Panel. <https://doi.org/10.6084/m9.figshare.21701624.v2>
- DeSanto, D., & Nichols, A. (2017). Scholarly metrics baseline: A survey of faculty knowledge, use and opinion about scholarly metrics. *College & Research Libraries*, 78(2), 150-170. <https://doi.org/10.5860/crl.78.2.150>
- Ferrier-Watson, A. (2019). *Traditional metrics, altmetrics and researcher profiles: A survey of faculty perceptions and use* [Research report, Victoria University of Wellington]. <http://researcharchive.vuw.ac.nz/handle/10063/8362>
- Graf, L., Wendler, W. S., Stumpf-Wollersheim, J., & Welp, I. M. (2019). Wanting more, getting less: Gaming performance measurement as a form of deviant workplace behavior. *Journal of Business Ethics*, 157(3), 753-773. <https://doi.org/10.1007/s10551-017-3688-y>
- Hammarfelt, B., & Haddow, G. (2018). Conflicting measures and values: How humanities scholars in Australia and Sweden use and react to bibliometric indicators. *Journal of the Association for Information Science and Technology*, 69(7), 924-935. <https://doi.org/10.1002/asi.24043>
- Harzing, A. W., & Alakangas, S. (2016). Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison. *Scientometrics*, 106(2), 787-804. <https://doi.org/10.1007/s11192-015-1798-9>
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature* 520, 429–431. <https://doi.org/10.1038/520429a>
- Hobson, J., & Hall, M. (2010). Publish or perish?: The challenge of journal rankings for Māori-related research. *MAI Review*, (3), 1-5. <http://www.review.mai.ac.nz/mrindex/MR/article/view/408/552.html>
- HuMetricsHSS. (n.d.). *HuMetricsHSS: Humane metrics initiative*. <https://humetricshss.org/>
- Karanatsiou, D., Misirlis, N., Vlachopoulou, M. (2017). Bibliometrics and altmetrics literature review: Performance indicators and comparison analysis. *Performance Measurement and Metrics*, 18(1), 16-27. <https://doi.org/10.1108/PMM-08-2016-0036>
- Kidman, J., Chu, C., Fernandez, S., & Abella, I. (2015). *Māori scholars and the university* (13RF01). Ngā Pae o te Māramatanga. <https://www.maramatanga.ac.nz/sites/default/files/project-reports/Kidman%20Chu%20Fernandez%20Abella-%20Maori%20Scholars%20Final%20report%202015.pdf>
- Kondakci, Y., Zayim-Kurtay, M., Kaya-Kasikci, S., Senay, H. H., & Kulakoglu, B. (2021). 'Scaling' the academia: Perspectives of academics on the impact of their practices. *Research Evaluation*, 30(3), 370-381. <https://doi.org/10.1093/reseval/rvab015>
- Locke, K., & Bensky, L. (2022). Invisible women: The PBRF as a regime of inequality in Aotearoa New Zealand. *Knowledge Cultures*, 100(2), 52-67. <https://doi.org/10.22381/kc10220223>
- Magnus, D. (2013). Overthrowing the tyranny of the Journal Impact Factor. *The American Journal of Bioethics*, 13(7), 1-2. <https://doi.org/10.1080/15265161.2013.807709>
- McAllister, T. G., Kidman, J., Rowley, O., & Theodore, R. F. (2019). Why isn't my professor Māori? A snapshot of the academic workforce in New Zealand universities. *MAI Journal*, 8(2), 235-240. <https://doi.org/10.20507/MAIJournal.2019.8.2.10>

- McAllister, T. G., Kokaua, J., Naepi, S., Kidman, J., & Theodore, R. (2020). Glass ceilings in New Zealand universities: Inequities in Māori and Pacific promotions and earnings. *MAI Journal*, 9(3), 272-285. <https://doi.org/10.20507/MAIJournal.2020.9.3.8>
- Ministry of Business Innovation and Employment. (2022). *Te Ara Paerangi: Future pathways white paper 2022*. <https://www.mbie.govt.nz/assets/te-ara-paerangi-future-pathways-white-paper-2022.pdf>
- Naepi, S. (2021). Shifting tides: The potential and limitations of critical mass to change discipline research habits. *Journal of New Zealand Studies*, (33), 35-35. <https://doi.org/10.26686/jnzs.iNS33.7380>
- Olive, R., Townsend, S., & Phillips, M. G. (2022). 'Not everything that can be counted counts, and not everything that counts can be counted': Searching for the value of metrics and altmetrics in sociology of sport journals. *International Review for the Sociology of Sport*, 58(3), 431-454. <https://doi.org/10.1177/10126902221107467>
- Onyancha, O. B. (2022). Indigenous knowledge, traditional knowledge and local knowledge: What is the difference? An informetrics perspective. *Global Knowledge, Memory and Communication*. Advanced online publication. <https://doi.org/10.1108/GKMC-01-2022-0011>
- Phillips, M. G. (2020). Sizing up sport history journals: Metrics, sport humanities, and history. *The International Journal of Sport History*, 37(8), 692-704. <https://doi.org/10.1080/09523367.2020.1796652>
- Price, R. (2022) *Results from the 2021 responsible metrics state of the art survey*. The Bibliomagician. <https://thebibliomagician.wordpress.com/2022/08/04/results-from-the-2021-responsible-metrics-state-of-the-art-survey/>
- Roa, T., Beggs, J. R., Williams, J., & Moller, H. (2009). New Zealand's Performance Based Research Funding (PBRF) model undermines Maori research. *Journal of the Royal Society of New Zealand*, 39(4), 233-238. <https://doi.org/10.1080/03014220909510587>
- Royal Society London. (2023). *Research culture*. <https://royalsociety.org/topics-policy/projects/research-culture/>
- Salesa, J. (2020). *Report back on the performance-based research fund review*. Ministry of Education. <https://www.education.govt.nz/assets/Documents/Further-education/PBRF-Review/CAB-20-MIN-0352-Report-of-the-Cabinet-Social-Wellbeing-Committee-Minute-....pdf>
- Sarpong, J. (2021). *Research autonomy and "marketisation" in higher education: A case study of two New Zealand universities [Doctoral thesis, University of Auckland]*. University of Auckland ResearchSpace. <https://researchspace.auckland.ac.nz/handle/2292/57954>
- Smith, L. T., Larner, W., Phipps, D., Town, I., Underhill-Sem, Y., Williams, M., & Mischewski, B. (2020). *E koekoe te tūi, e ketekete te kākā, e kūkū te kererū: Toward the tertiary research excellence evaluation (TREE)*. Ministry of Education. <https://assets.education.govt.nz/public/Documents/Further-education/PBRF-Review/The-Report-of-the-PBRF-Review-panel-E-koekoe-te-tuie-ketekete-te-kaka...pdf>
- Tawhai, V., Pihera, K., & Bruce-Ferguson, P. (2004). Does the PBRF need reshaping? A new Maori educational institution's perspective. *Research and development in higher education: Transforming knowledge into wisdom holistic approaches to teaching and learning*, 27, 333-340. <https://www.herdsa.org.au/publications/conference-proceedings/research-and-development-higher-education-transforming-32>
- Tilbury, C., Bigby, C., & Hughes, M. (2022). The production and dissemination of Australian social work scholarship: A citation analysis. *Australian Social Work*, 75(4), 407-419. <https://doi.org/10.1080/0312407X.2020.1798482>
- University of Bath. (2017). *Principles of research assessment and management*. <https://www.bath.ac.uk/corporate-information/principles-of-research-assessment-and-management/>

- Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., Viney, I., Wouters, P., Hill, J., & Johnson, B. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. <https://www.ukri.org/wp-content/uploads/2021/12/RE-151221-TheMetricTideFullReport2015.pdf>
- Yang, S., Zheng, M., Yu, Y., & Wolfram, D. (2021). Are Altmetric.com scores effective for research impact evaluation in the social sciences and humanities? *Journal of Informetrics*, 15(1), 101120. <https://doi.org/10.1016/j.joi.2020.101120>



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

