

Open Science, open issues

Sarita Albagli
Maria Lucia Maciel
Alexandre Hannud Abdo
Editors



4

The road less travelled: optimizing for the unknown and unexpected impacts of research

Cameron Neylon

*Two roads diverged in a wood, and I—
I took the one less traveled by,
And that has made all the difference.*

Robert Frost – from *The Road Not Taken*

THE PURPOSE OF THE ACADEMY

What is research for? What purpose does the academy, or the university, or the research institute, serve? These are questions we shy away from, both because it is difficult to reach consensus but also because it requires a level of self-examination that is uncomfortable. Probing our own motivations and the motivations of those who fund us can be unsettling.

There may be broad societal agreement that research is a generally good thing, but there is very little agreement on why that might be. Governments with a market orientation see an economic value in innovation. Campaigners look to academic experts to question the government's focus on markets. Patients and their families hope for new treatments, environmentalists may look for studies that show the damage that the factories

producing those treatments can cause. Technologists might point to the value of science in helping us to understand and tame the natural world. Humanists point to the value of the humanities in helping us to understand ourselves so as to obviate the need to tame the outside world.

It may be difficult to reach agreement on what research should deliver. These are deep questions of values. But we should be able to reach agreement that there is a responsibility on the part of the academy to those who pay for research. That responsibility is to deliver well in accordance with those values. Delivering well might mean efficiency or it might mean effectiveness. Even that is not clear. But delivering well is a responsibility we should assume and talk about.

TESTING THE PERFORMANCE OF INSTITUTIONS

We build institutions to carry out research. In a perfect world we would build these institutions based on clearly articulated shared values. We would use those values to craft an effective and useful statement of the mission of the institution and we would then assess performance against that mission.

Mission statements can be woolly aspirational statements, but the best are useful strategic decision making tools. Crafting a good mission statement is a substantial challenge. In particular crafting statements that assist in making objective performance assessments, while still reflecting the full set of shared values, is at best difficult and often impossible. Easily measured and instrumental goals support instrumental assessment, which drives instrumental behavior – performing against the measure, rather than performing against the mission.

In practice, we measure what we can, and that in turn becomes the *de facto* mission. Rather than asking whether we are doing well at generating new knowledge and how effectively we are

transferring it to those who can use it we ask how many articles have been published and what journals they were published in. The problems of instrumentalism and naïve metrication of research assessment are well rehearsed. We will not spend significant time on it here but those criticisms should be borne in mind.

INSTRUMENTAL STRATEGIES AND INSTITUTIONAL LEADERSHIP

The issue with metrics is not the metrics themselves. These proxies or indicators measure what they measure. The problem arises when strategic decisions are made on the basis of the measures themselves, rather than assessment of performance against a well articulated mission. The problem is not that number of articles, or H-index, or grant income can't answer a question; it is that they cannot provide complete answers to the questions that *should* be asked – how *productive* is a researcher, what is their *influence* in the community, in what ways do they *contribute* to the institution.

These questions, and others that would follow from a well designed mission statement, are not straight forward to answer. They will not be addressed by any single indicator, nor any simple “basket of metrics”. Indicators and metrics can only ever be data to support the strategic decision making. Too often, in a search for an illusory objectivity we reach for quantitative measures as a way of avoiding the responsibility to make those decisions.

The unique contribution of an effective institutional leader will be to make informed strategic decisions. Reliance on league tables and metrics, the posting of thresholds or performance targets is a sign of a lack of confidence in their ability to make those decisions. The best institutional leaders will use metrics and other indicators as data to assist in making decisions. They will not use quantitative measure *to* make decisions. They will have a diverse set of data at their fingertips and an understanding of how to integrate that to assess a wide diversity of research activities.

DIVERSE PORTFOLIOS AND DELIVERING ON MISSION

The future impact of research is unpredictable and investment in research is risky. The rational response to this is to hold a diverse portfolio. At each level of granularity; investigator, group, department, institution it makes sense to have a range of different activities that as a collection optimize the opportunity for delivering value.

The true cost of the instrumentalism described above has been homogenization. Institutions are all trying to climb the same league tables based on narrow criteria. Only a very small number of highly prestigious institutions have the self-confidence to carve out their own path. The irony is that institutions worldwide seek to rise up league tables so as to be like Harvard or Stanford or Cambridge, while those institutions do what they do largely because they ignore those same tables.

In the end what research is for is a question for institutions, communities, nations and global publics to answer for themselves. But when institutions address that question they should focus more on what makes them unique instead of what makes them a pale imitation of Princeton or Oxford. Diversity of mission and focus at the institutional level will aid in delivering on mission at the national and global levels by creating a portfolio of institutional profiles.

Diversity at the institutional level will also provide space for a more diverse range of researchers generating more diverse outputs and more diverse impacts. Clearly this creates challenges for institutional leaders and effective institutional leadership. Ultimately the challenge for assessment is developing a sufficiently diverse set of indicators to support the tracking and management of such a diverse portfolio. Perhaps even more challenging is to know how to combine those indicators to support effective decision-making.

IMPACTS AND INDICATORS

While we recognize that agreeing on the values and mission for the research enterprise is challenging it will nonetheless be useful to consider the different classes of results we might wish for and the extent to which they can be measured. “Impact” may be a dirty word in many research circles but it is nonetheless a useful technical term.

There are a range of different definitions in use, but in the current context I will use a meaning that expands on that used by the Australian Research Council¹, Research Councils UK² but including the scope described by the LSE Impact of Social Sciences project³: the change in the world that results from the dissemination of research outputs. We can speak of different forms of impact, including economic impacts such as job creation but certainly not stopping there. We can also consider impacts in the areas of policy, education, culture, environment and health. I explicitly include impacts on research activities as well as “wider impacts”.

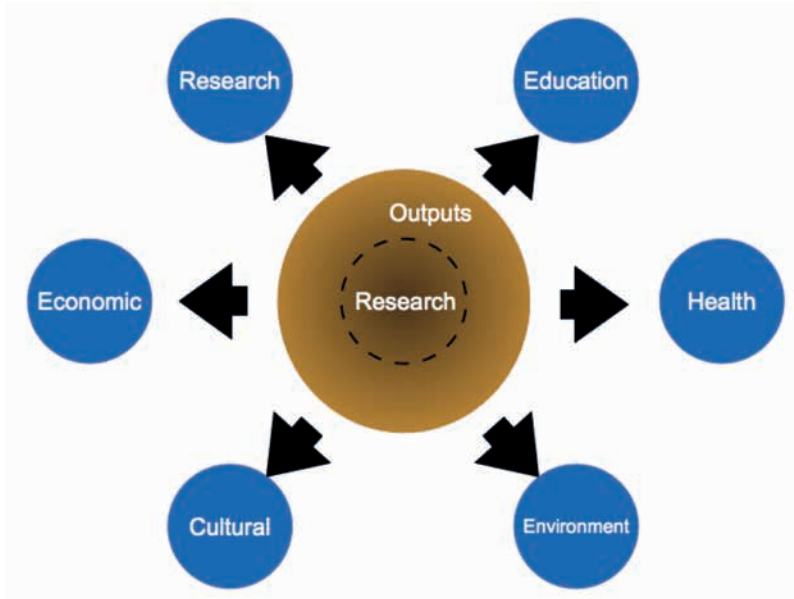
Again the prioritization of different classes of impacts is a matter for community discussion but we can recognize that these impacts depend on the outputs of research being disseminated to those places where they can be applied. The outputs might be concepts, skills, new technologies or approaches, or they may be people. There will be diversity in outputs, impacts and the paths that join them together.

¹ Available on: <http://www.arc.gov.au/research-impact-principles-and-framework#Definition>. Access on: June 21, 2015

² Available on: <http://www.esrc.ac.uk/funding-and-guidance/impact-toolkit/what-how-and-why/what-is-research-impact.aspx> . Access on: June 21, 2015

³ Available on: <http://blogs.lse.ac.uk/impactofsocialsciences/introduction/> . Access on: June 21, 2015

Figure 1. Varying forms of research impact. These different categories of effect have little in common being linked only by the process of research which leads to them. Research is transmitted through outputs (although that distinction is becoming more porous, shown by the dotted line) and on through some process into outcomes and impacts.



PROXIES, INDICATORS AND THEIR MEANING

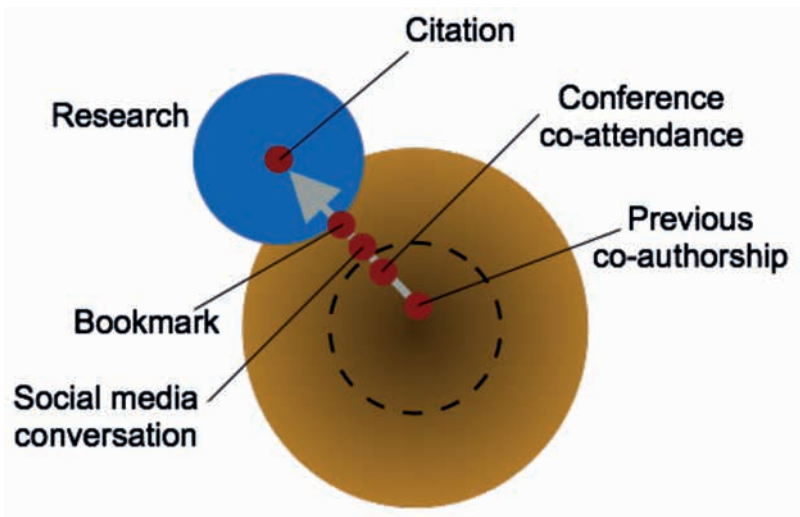
Impacts are what we ultimately seek to maximize, but in practice they can almost never be measured directly. Outputs by contrast tend to be easier to track and measure. Our traditional focus on research articles and their citation is driven at least in part by the ease of tracking and quantifying their number.

If we focus research impact as an example, our aim is to measure the change in future research that results from a given project, a

given output, or the work of a given researcher. What we have traditionally measured is productivity in outputs and citations. For all the potential diversity of possible impacts and outputs our view has been restricted almost entirely to these two sets of proxies. It is not only that metrication and instrumentalism are problematic in and of themselves but that our field of view has been horribly limited.

The movement of research online and greater general availability of information about the research enterprise has provided new proxies that have the potential to provide a richer view (NEYLON; WU, 2009; PRIEM et al. 2010). We are able to track discussions and use of research in a much wider range of places, from social to mainstream media, through bookmarking services to secondary sources like Wikipedia and policy documents.

Figure 2. The growing set of indicators and proxies that might be useful in measuring pathways to the impact of research on further research. In the past we only had the publication and citation events to work with.



This richer variety of data has the potential to provide a much more diverse view of the flow of knowledge and to support assessment of a wider diversity of activities. At the same time there are often questions raised as to what these new measures *mean*. Can a tweet tell us as much as a citation. Do bookmarks in Mendeley really mean someone has read an article? Is it necessarily the case that mainstream media coverage means the research is good or useful or important?

COUNTING PROXIES OR TELLING STORIES

An objection often raised for any quantitative measure, including citations is that counting is misleading. Often this appears in the form of a statement such as “citations can also be negative” or “popularity is not a sign of impact”. These criticisms become stronger when we look at proxies such as downloads or social media mentions, where the numbers can be large and where popularity (appears as though it) might play a stronger role.

A more productive way to use these proxies can be to use them to discover and tell stories. With social media in particular the numbers can be misleading due to reinforcement effects. It is important to investigate who is talking about a given research output and what they are saying (as well as who to).

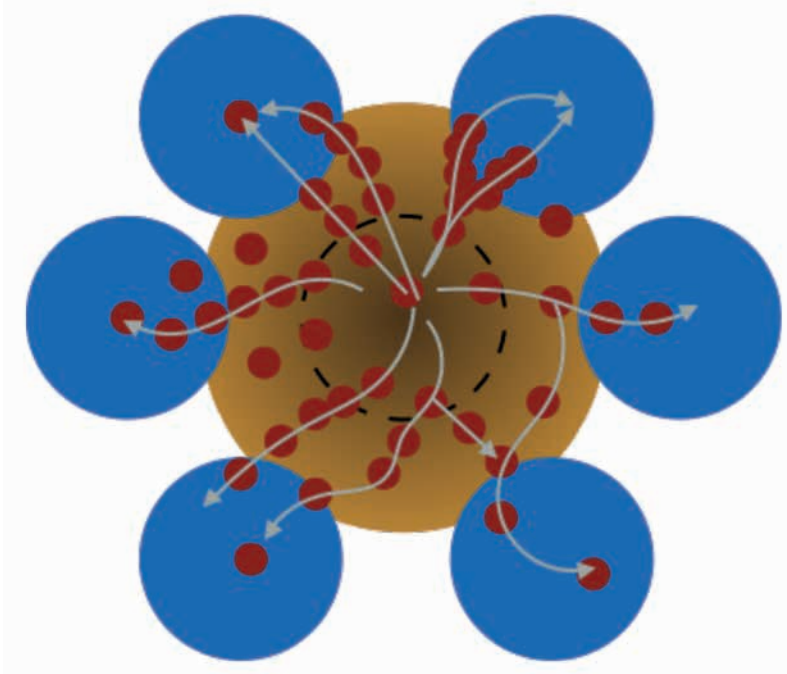
For example a story I often tell relates to Twitter. I was investigating papers published by the University of Cape Town with PLOS using data from the PLOS Article Level Metrics service and also the altmetric.com service, which provides information on the geolocation of tweets. There were very few tweets about this South African corpus of papers that originated in South Africa. However one paper (JEWKES et al. 2011) stood out as having some South African activity.

This was a paper on the relationship between HIV status and domestic violence. In particular the accounts talking about the paper were associated with women's crisis centres, sexual health clinics and support centres for minority sexual orientations. Furthermore I could identify the specific account and therefore the person that was disseminating this research to places where it might be directly applied. The counting of tweets was not very useful here, but identifying who was behind those tweets told a powerful story.

However there is a sense in which both the objection to quantitative metrics and the ability to tell stories expose a basic fallacy in the way we think about metrics old and new. Throughout this text I have been careful to refer to measures as "proxies" or "indicators". Too often the objections arise about either "meaning" or quantitation because of an assumption that a metric *itself* is what matters. Of course this is never the case, what matters is not citations but the influence they are a proxy for, not social media mentions but the way they inform us about communities using the research, not downloads but the usage that they signals.

All of these measures are merely proxies for things that we care about, but in many cases they are not even that. They are indicators of the flow of knowledge. It is more useful perhaps to think of them as flares that light up when knowledge flows past a point on a path. The same flares may result from many different pathways, and knowledge may flow through many different pathways to the same destination.

Figure 3. A figurative image of pathways of knowledge transfer (arrows) and the signals that arise (red dots). The observable word is only the red dots and the majority of our research assessment systems are based on only two of those dots, citations and publication events.



We can think of the pathway to impact as a set of knowledge flows, where the flow itself is invisible. All we have are indicators that signal parts of that flow. It becomes clear that it is only by combining multiple measures that we can pick out a specific path. The paucity of our traditional measures also becomes clear, one or two lights blinking on one single (assumed) pathway tells us little or nothing that is useful. Finally the question of “what does this metric means” falls away. The metric doesn’t *mean* anything in isolation, it is an indicator, mere data that in combination with

other data may help us to understand the pathways through which a given piece of knowledge is disseminated.

We can also integrate the narrative view with a quantitative view. The stories are instances of knowledge flow down various pathways. Quantitative analysis of indicators can help us to understand the overall flows and their paths as well as helping us to identify specific instances of that flow. The story above is simply a very crude example of this form of analysis but more sophisticated approaches are certainly possible.

THE ROAD LESS TRAVELLED, THE ROAD UNKNOWN

This pathway model is potentially very powerful. Firstly it helps us to avoid the fallacy that a countable metric is itself what matters. Secondly it provides a route into more sophisticated analytical approaches that do not assume prior knowledge of what the pathways are. This brings the potential of “big data” analytics to bear on the problem of identifying and mapping the pathways.

This is superficially similar to many other models of how research leads to impacts. Most models describe, or aim to surface, some form of pathway or pathways. The Becker Model of Impact for biomedical sciences focuses on specific forms of impact and identifies indicators that lie on the path towards each of them (BERNARD BECKER MEDICAL LIBRARY, s/d). The Payback model and BRIDE tool developed from it (SCOBLE et al., 2010) have a similar conceptual framework. Many of these models build on diffusion of innovation theory, in itself a linear model (WALTER et al., 2013).

Even in those approaches where participants and stakeholders are engaged in defining desired impacts such as the Participatory Impact Pathways Analysis approach (STEPS CENTER, s/d) the focus

is on defining *the* pathways that exist, or are desirable, for further monitoring. Overall, existing models and methodologies assume that there are known (or discoverable) and generally linear pathways through which knowledge or insight flows to create impacts.

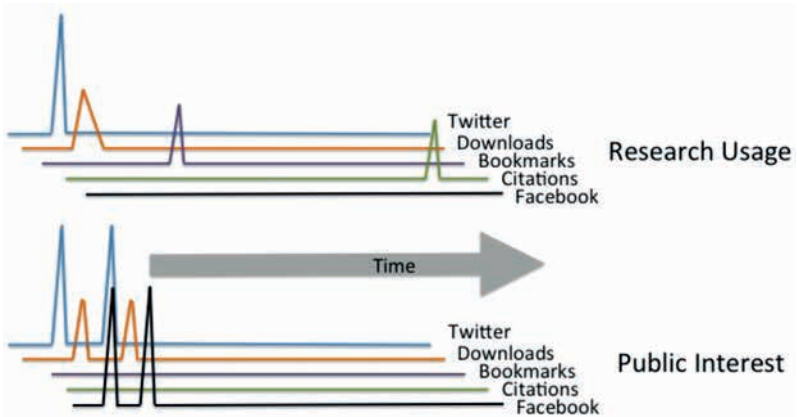
This leads to assessment frameworks in which various indicators are tied to specific impacts, and therefore specific pathways. They develop matrix approaches in which, by measuring the presence of specific indicators, sometimes through quantitative approaches, sometimes qualitative, evidence of specific impacts (or their future potential) is provided. In turn the same matrices can be used to optimize research dissemination so as to maximize those desired impacts.

A “HIDDEN PATHWAYS” MODEL OF KNOWLEDGE FLOWS TO RESEARCH IMPACTS

In contrast to these linear and explicit models, the model developed in the previous section assumes that the pathways are unknown and probably highly branched. There is an implicit focus on more granular indicators and to some extent to more quantifiable ones, as opposed to qualitative and narrative indicators or evidence. Finally there is an implicit requirement that indicators can be tied to *events*, that is, they can be fixed in time.

Formally, this hidden pathway model is described as a set of measurable channels (indicators) in which signals can be measured. These signals are indicative of processes (knowledge flows along defined pathways). The signals can be thought of as flares or blinking lights that show when some particular knowledge transfer is occurring along a pathway. Any given indicator may be attached to none, one, or many pathways. The pathways themselves are not observable, but can only be inferred.

Figure 4. Signal patterns from different underlying processes. Both processes lead to signals from twitter and download channels but a research use also shows later bookmarking and citation activity. Public interest shows greater correlation with Facebook and tighter time domain correlation. The signal patterns are hypothetical based on non-quantitative observation of specific data sets.



The means of inferring a pathway is through identifying patterns of signal activities that occur across sets of indicators (signal channels). For instance a hypothetical “scholarly knowledge transfer” pathway might involve a tweet (through which a scholar discovers a work), a download or view (reading the work), bookmarking and then citation in the formal literature. Impact on a patient group might start from the same place (a tweet, a download) and then branch off through a Facebook conversation and on to wall posts on a patient-focused service.

Of course, all these processes are occurring at the same time, leading to very complex signal patterns, which need to be disentangled. Broadly speaking this means using maximum likelihood methods to model the probability distributions of sets of possible processes that explain the observed patterns of

signals. Essentially the aim is to embrace the larger sets of data we have available to us so as to cast the problem as one of time domain signal processing. There are strong analytical methods from engineering and other disciplines that are designed to tackle precisely this class of problems; trying to untangle the multiple underlying processes that are giving rise to a complex multichannel signal.

MAPPING UNKNOWN PATHWAYS

The advantage of this conceptual approach is that it creates the potential not just to identify flows down the pathways we (think we) know about but also to surface new pathways. Instead of either assuming specific pathways exist, or seeking to surface them through conversations with stakeholders, it makes it possible to start from an assumption that there are knowledge flows that no-one is aware of and make an attempt to discover them.

Whether this is possible in practice is uncertain. Such approaches require large quantities of data with high quality time information. While the quantity of data we have is certainly increasing it is not clear that it is sufficient to surface unknown pathways. Even if we have the data the quality of the time information is generally rather poor. For tweets we can utilize a time stamp, but article download data has very variable time resolution, and is also collected differently by different organizations. In attempting this form of analysis we should identify the weaknesses in our data.

A side effect of analysis that seeks to identify the underlying processes occurring is that there is also the potential to detect signals that arise from processes not related to desired impacts. Such signals might include errors or problems in data collection or processing. Or they might reflect attempts to game metrics. We are already aware for instance that strong signals in a single channel (such as downloads) that do not correlate with signals

in other channels (such as bookmarks or tweets) are indicative of gaming.

There are three broad weaknesses with this approach. The first, discussed above is the dependency on data scale and quality. In practice we may only be able to distinguish the strongest signal correlations and therefore not achieve the insight into the unknown pathways that we would desire. The second is that it is clear that the pathways themselves are rapidly changing at the moment. This complicates the analysis, and although not rendering it impossible, puts further demands on data scale and quality to obtain new insights. In an ideal world the best way in would be to have a set of data in a stable environment. The irony of course it that we are interested in the analysis precisely because the environment is not stable.

The final weakness is the most central. These approaches can not identify the actual pathways. We can only ever infer that a pathway we can qualitatively describe corresponds to a probabilistic model of signal correlations. More generally such an analysis can not provide direct evidence of impact itself. The signals indicate underlying processes, not change in the word. To use this analysis to help us understand or optimize impact we need to embed it in a social practice, which leads us back to the need for articulating values.

RESPONSIBLE MANAGEMENT AND ARCHITECTING OF THE RESEARCH ENTERPRISE

The focus of this hidden pathway model of knowledge transfer is to exploit a technical analytical capacity to better understand and optimize the pathways the lead to research impacts. It is fundamentally technological. Yet I started with what is fundamentally a social issue of responsibility and values. How do we bring these together?

The link for me is through leadership, management and institutional design. In practice the conversation about what the shared values for the research enterprise is an ongoing one. These values will evolve and change as communities' needs change and as our capacity to address them changes. I make an assumption that a way to address both this issue of change and uncertainty as well as the unpredictability of research outcomes is through embracing diversity at a range of levels. Diversity of goals, of skills, of outputs and of research agendas allows, potentially, for buffering of capacities as well as agility in response to changing needs, as well as providing many controls that can be tuned to optimize impacts.

The defining characteristic of research is its unpredictability. If we knew the answer we wouldn't need to do the research. Picking winners is near impossible. This makes it imperative that we design our institutions at the systems level. Decisions about individual projects, or appointments, or modes of dissemination will always be informed guesses. But we can tune the processes by which we make those decisions so as to optimize the *average* outcome. It is entirely possible to design an electrical circuit without needing to know what path an individual electron will take.

A central design challenge for such systems is to optimize for the possibility of unexpected outcomes and impacts, unexpected pathways to impact. It is a matter of faith amongst researchers that the most important insights arise from serendipity. Yet we focus almost exclusively on known modes of communication to specific, known audiences defined by specific journals. In truth we *do not even know* how much research impact arises in the expected places versus the unexpected. We try to measure *expected* impact (or more strictly progress towards it) through a horrendously narrow, albeit expanding, set of proxies that are totally inadequate to the task, yet I argue we have a responsibility to also seek to maximize the *unexpected*.

The hidden pathways model I have described here seeks to address the lack of data that should trouble a responsible

institutional leader. But in addressing that issue it also takes away any comfort that can be derived from the measurement of progress against naïve and simplistic rankings that currently characterize institutional decision making.

Such simple rankings are comfortably normative; higher is “good”, downward is “bad”. Everyone agrees, even those who are violently opposed to the rankings themselves. In a model focused on pathways to diverse impacts there is no “up” or “down”, there is no normative position on which impacts are better or more important. These are not even decisions that leaders can themselves take, involving as they do whole communities.

The responsibility for leaders therefore becomes greater, and in many senses the freedom to act becomes less. A leader is a curator of the conversations that articulate these values, the guardian and caretaker for a useful mission statement, and an engineer who must constantly seek to tweak a thousand settings to optimize importance.

This is perhaps not the skillset that characterizes today’s generation of institutional leaders – it is however one that aligns closely with successful managers of online communities. This may require a generational change, but in turn the institutions of our future will be the ones that are successful in a world of online communication. It may indeed be the road less travelled, but with luck it will make all the difference for the future of a successful, community embedded and responsible research enterprise.

REFERENCES

AUSTRALIAN RESEARCH COUNCIL. *Research impact principles and framework*. [online]. Available on: <http://www.arc.gov.au/research-impact-principles-and-framework#Definition>. Access on: June 20, 2015

BERNARD BECKER MEDICAL LIBRARY. *Assessing the impact of research – How to use the model*. [online]. Available on: <https://becker.wustl.edu/impact-assessment/how-to-use>. Access on: June 20, 2015

ECONOMIC AND SOCIAL RESEARCH COUNCIL (UK). *What is research impact*. [online]. Available on: <http://www.esrc.ac.uk/funding-and-guidance/impact-toolkit/what-howand-why/what-is-research-impact.asp>. Access on: June 20, 2015

FROST, Robert. The Road Not Taken. In: FROST, Robert. *Mountain Interval. Selected Poems*, s/l, s/ed: 1920. Available on: <http://www.poets.org/poetsorg/poem/road-not-taken>. Access on: June 21, 2015

JEWKES, R.; SIKWEYIYA, Y.; MORRELL, R.; DUNKLE, K. The relationship between intimate partner violence, Rape and HIV amongst South African Men: a cross-sectional study. *PLoS One*, v.6, n.9, p. e24256, 2011. doi:10.1371/journal.pone.0024256.

LONDON SCHOOL OF ECONOMICS. *The impact blog – introduction*. [online]. Available on: <http://blogs.lse.ac.uk/impactofsocialsciences/introduction/>. Access on: June 20, 2015

NEYLON, C.; WU, S. Article-level metrics and the evolution of scientific impact. *PLoS Biol*, v.7, n.11, p. e1000242, 2009. doi:10.1371/journal.pbio.1000242.

PRIEM, J.; TARABORELLI, D.; GROTH, P.; NEYLON, C. *Altmetrics: a manifesto*. [online]. [S.l.][s.n.] 2010. Available on: <http://altmetrics.org/manifesto/>. Access on: June 20, 2015

SCOBLE, R.; DICKSON, K.; FISHER, J.; HANNEY, S. *Research impact evaluation, a wider context: findings from a research impact pilot*. [online]. [S.l.][s.n.] 2010. Available on: <http://bura.brunel.ac.uk/handle/2438/4149>. Access on: June 20, 2015

STEPS CENTRE. *Participatory Impact Pathways Analysis (PIPA)*. [online]. Available on: <http://steps-centre.org/methods/pathways-methods/vignettes/pipa/>. Access on: June 20, 2015

WALTER, I.; NUTLEY, S.; DAVIES, H. *Research impact: a cross sector review - literature review*. [online]. [S.l.][s.n.] 2003. Available on: <http://www.ruru.ac.uk/pdf/LSDA%20literature%20review%20final.pdf>. Access on: June 20, 2015