

Celebrating women conducting research in freshwater ecology ... and how the citation game is damaging them

Barbara J. Downes^{A,B} and Jill Lancaster^A

^ASchool of Geography, The University of Melbourne, 221 Bouverie Street, Parkville, Vic. 3010, Australia.

^BCorresponding author. Email: barbarad@unimelb.edu.au

Abstract. We highlight women's contributions to freshwater ecology by firstly considering the historical context and gender-based barriers faced by women attempting to gain an education and secure research jobs in science over the past 100+ years. The stories of four remarkable, pioneering women in freshwater ecology (Kathleen Carpenter, Ann Chapman, Rosemary Lowe-McConnell and Ruth Patrick) illustrate the impact of barriers, emphasise the significance of their contributions and provide inspiration for the challenges ahead. Women still face barriers to participation in science, and the second part of the paper focuses on a current form of discrimination, which is citation metrics used to measure the 'quality' or 'impact' of research. We show that arguments that citation metrics reflect research quality are logically flawed, and that women are directly disadvantaged by this practice. Women are also indirectly disadvantaged in ecology because they are more likely to carry out empirical than theoretical research, and publications are generated more slowly from empirical research. Surveys of citation patterns in ecology reveal also that women are less likely to be authors of review papers, which receive three times more citations than do original articles. Unless unfettered use of citation metrics is stopped, research will be damaged, and women will be prominent casualties.

Additional keywords: citation counts, gender bias, h-index, journal impact factor, research impact, research quality.

Received 11 November 2018, accepted 17 January 2019, published online 16 April 2019

Introduction

The invitation to write an opinion piece for this special issue, which aims to highlight women's contributions to freshwater science, was accompanied by the dilemma of identifying a suitable thesis that would also be of general interest. We have chosen to look at women's contributions initially in a historical context, and particularly in the context of the barriers that have been erected over the years to prevent women participating in science at all. Our motivation arises from the simple concept that the past causes the present and so the future. Therefore, we are better prepared for the future if we understand the past. Analogously, freshwater science progresses by understanding the achievements of past researchers (often in diverse disciplines), the sequence of developments that have led to the current state of knowledge, and this helps identify potentially fruitful ways forward.

The past difficulties and injustices endured by women in science can be maddening to read. Considering the stories of some women freshwater ecologists who weathered these past adversities, nevertheless, provides both illustration and inspiration for the current and future challenges. As we will argue, these difficulties are far from over and women in science continue to be disadvantaged by gender-biased practices. One major difficulty now faced by *all* researchers in all disciplines is the inappropriate and unacceptable use of citation metrics to

evaluate the 'quality' of a research contribution or of a researcher. However, as we will show, women are disproportionately disadvantaged by this practice, which threatens their ability to secure employment, promotion and research funds. This is alarming. So, in this paper we hope to highlight and therefore celebrate some past contributions of women to freshwater ecology, but also to identify a major challenge that women face right now. An awareness and understanding of this history will, we hope, be of benefit to the current battle in the war against sexual discrimination.

This paper has two main parts. The first part discusses the historical barriers faced by women who were interested in learning or having a career in science. This included getting an education, getting a qualification, getting a job, keeping a job, and the freedom to pursue any research question. We illustrate how these barriers affected the careers of four pioneering women in freshwater ecology and highlight some of the significant contributions these women made, despite the odds. These historical barriers directly discriminated against women, practices that are now illegal in many nations. However, women still face barriers to participation in freshwater ecology. The second part of this paper focuses on citation metrics; these are logically flawed, which we will explicate. Using examples that are relevant to freshwater ecology, we use surveys of publication and citation patterns to demonstrate that women pay a heavier

price when citation metrics are used to measure alleged research quality or impact. For freshwater ecology, or indeed any discipline to thrive, this logically flawed practice must stop. We finish with some suggestions of how that might be achieved.

The historical context and previous barriers for women in science

Contributions of women to *any* science discipline were scarce pre-1900, because women were deliberately excluded from tertiary education and many kinds of employment. In this section, we briefly describe the sequence of barriers that led to this exclusion, and the changes that allowed women to participate. This history is well known and there are many learned publications; some events may be familiar to older readers, but perhaps less so for younger generations. Nevertheless, the legacy of these circumstances is with us today, and an awareness of the historical context is valuable to understanding the present barriers that women face.

The reasons men concocted for barring women from education and employment would be hilarious, if they had not been so devastatingly effective. The initial arguments were simply that women were inherently inferior and incapable of any intellectual pursuit. Towards the end of the 19th century, allegedly scientific arguments about the inferiority of women's brains were no longer widely acceptable, so the new arguments focussed on why education was detrimental to women's health (Hubbard 1990). In 1874, Edward H. Clarke published a book full of alleged medical explanations why women should not be educated – this was a highly successful book that went through 17 editions. The main thrust of his argument was that education would prevent the normal development of women's reproductive physiology and they would become sterile, and likely mentally unhinged as well. Women were meant to be nothing more than a means to care for and produce more men. Because women are physiologically different from men, they must, therefore, be abnormal and all things related to reproduction were considered disabilities (Clarke 1874). As Hubbard (1990) reminds us, all the sciences (e.g. biology, psychology, physiology, medicine) are ultimately male constructs and alleged scientific arguments have long been used to disqualify women from participating in society.

Being an active scientist requires training (i.e. tertiary education) and an opportunity to practice (i.e. a job), and women were actively excluded from both until reasonably recently. A casual and selective perusal of university archives (and other sources) from institutions in the UK, USA, Europe and Australia makes for sobering reading. For brevity, we have not provided references for numerous historical facts that are well documented and easily verified, such as the dates that particular universities first awarded degrees to women. What follows is not an exhaustive history, but simply a selection of significant events that encapsulates the situation for women with academic interests generally (Fig. 1). The situation for women interested in science was probably worse. Our objective is not to complain about sexist practices, but to identify them and consider the legacy that remains.

In general, women were barred from university education until the 1870–1880s, regardless of when the university was

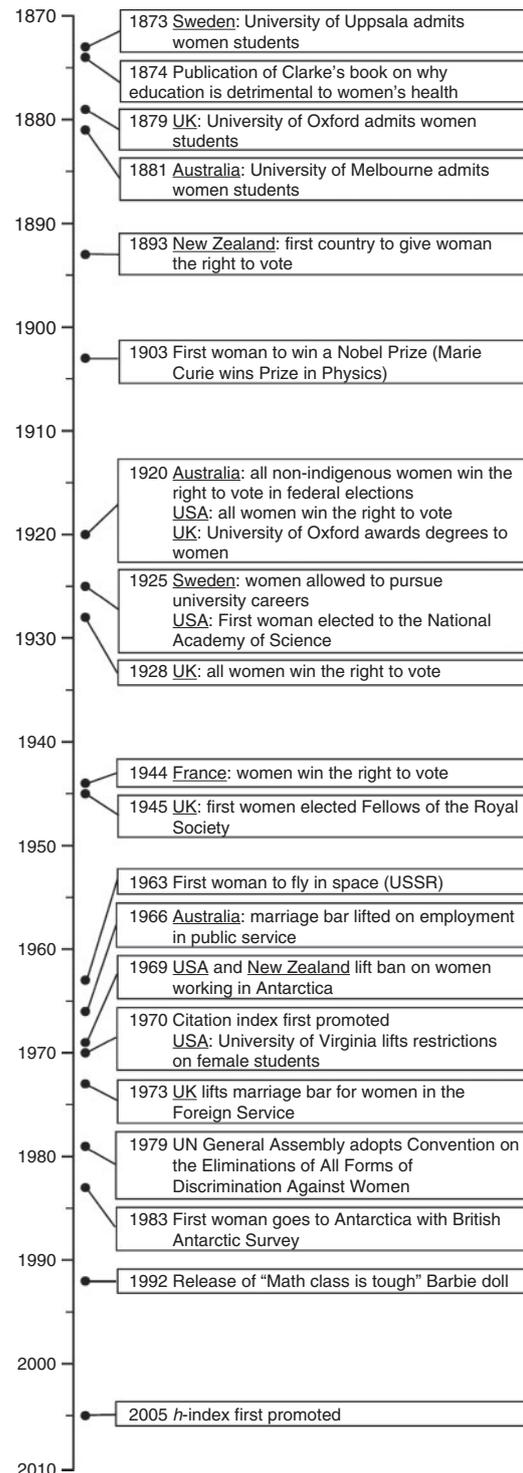


Fig. 1. Timeline showing a selection of notable events in the history of women in science. See text for further discussion.

founded. For example, the University of Oxford (UK) was founded sometime before 1167 but women were barred from attending classes for ~700 years, until the late 1870s; the University of Uppsala (Sweden), founded 1477, accepted

female students only ~400 years later in 1873; the University of Melbourne (Australia), founded 1853, accepted women only after 1881. Although this was extraordinary progress for the emancipation of women at that time, it often occurred only after many years of petitioning university administrators. Not all institutions fully embraced the idea and continued to create significant hurdles for women. For example, Columbia University (USA) allowed women to take exams from 1883, but initially barred them from attending classes with men so they had to study at home without the benefit of listening to lectures. Harvard University (USA) would not let women sit in classrooms with men and so created a separate, women-only college in 1879 (the 'Harvard Annex' which later became Radcliffe College), with lectures being delivered by male academics from Harvard University. By 1900, ~71% of colleges in the USA were co-educational; Virginia was the last state in the union to provide post-secondary co-education. The University of Virginia allowed women to register as 'special students' from 1892, but it was decades before these women could be awarded degrees and – not until 1972! – that the university lifted all restrictions on women.

Despite being allowed to attend classes in many institutions, being awarded degrees did not necessarily follow, which meant that women were still barred from graduating and being admitted formally into many universities. For example, the first women allowed to take, and who passed, the General Examination at the University of London (UK) in 1869 were awarded a 'Certificate of Proficiency' not a degree. Only a decade later, in 1880, did the University of London start awarding degrees to women. The University of Oxford awarded degrees to the first female graduates only in 1920, 50 years after they were allowed to attend lectures. Postgraduate degrees for women followed at many institutions by ~1900, but the numbers were few and even fewer in science disciplines.

Awarding academic positions to women had become reasonably common by the 1920–1930s, and a few full professorships were awarded by the 1940s, but proportional representation at all academic levels was poor and continues to be poor, especially in science disciplines (see section *Modern barriers to women in freshwater ecology*). Some women achieved teaching or tutoring positions years earlier, but they typically lacked the opportunities to carry out research. Because of the attitudes of men towards women (e.g. many still believed that women had inferior brains), many women in universities tried to emulate men and to disguise the fact that they were women, for example, they wore 'simple, unattractive clothes' (Patrick 1997, p. 5). Access to tertiary education had improved, but there were still significant hurdles to employment for women. In Sweden, for example, women could take university exams in 1873, but they could not pursue a university career or assume a higher position within the government until 52 years later in 1925. Barriers to women in securing academic positions not only stifled research by women, but also deprived students of female role models. As the biologist Ruth Hubbard said about the impact on her generation of Radcliffe students of not being taught by women: '...to study under Harvard's 'great men' ... was thereby denying us the realistic expectation that we might some day be equally great women.' (Hubbard 1990, p. 46).

Even after women secured opportunities to study and hold academic positions, sustained research contributions by women were not celebrated; for example, very few national academies of science elected women members before 1945 (Mason 1992), and women continued to be barred from whole research areas, such as anything to do with Antarctica. As usual, these exclusions were justified by flimsy excuses. For example, personnel at the British Antarctic Survey (BAS) sent a letter to a woman who applied in the 1960s that stated:

Women wouldn't like it in Antarctica as there are no shops and no hairdresser. [Jones 2012].

A more likely and puerile explanation is that

...the presence of women would wreck the illusion of the frontiersman – the illusion of [a man] being a hero. [quotation from Commander of US Antarctic operations; Chipman 1986, p. 87].

In 1983, the first woman went to Antarctica with the BAS, although women were still effectively barred from using UK bases and logistics in 1987 (Sudgen 1987). In the USA, it required action from the United States Congress in 1969 to finally lift the ban on American women working in Antarctica. The patently silly excuses for banning women from the continent of Antarctica in connection with research or research support are even more ridiculous when we recognise that women had carried out ship-based research in Antarctic waters, sailed with whaling ships, and had been visiting, staying and having children on sub-Antarctic islands for many decades before these bans on government-associated activities were finally lifted (Chipman 1986). At the other pole, Inuit women have lived in the Arctic for millennia. Fortunately, women are no longer banned from entire continents and are actively engaged in, or supporting, research on all parts of Antarctica, including its freshwater ecosystems.

Finally, it is worth noting that barriers for women getting jobs were not restricted to universities or research environments, and government legislation forbidding some forms of employment were pervasive throughout some societies. We refer particularly to marriage bars, which some countries implemented to restrict the employment of married women in general or in particular occupations, although they typically did not affect employment in low-paid and unskilled jobs. In Australia, married women were barred from teaching until 1956 and the bar on employment in the public service was lifted only in 1966 (Sawer 1996). Ironically, this occurred only after two women (one was a postgrad student, and both happened to be wives of university academics) chained themselves to the furniture in a hotel bar room in Brisbane to draw attention to the apparently unrelated issue of a Queensland law that prevented women being served alcohol in public bars (Lake 1999). In the UK, a marriage bar prohibiting married women from joining the civil service was not abolished until 1946 for the Home Civil Service and 1973 for the Foreign Service. However, it was only when the *Sex Discrimination and Equal Pay Acts* of 1975 was passed by the UK government that other employers could no longer force women to resign should they marry, as had been the custom at the British Geological Survey (Pennington 2015). As well as preventing married women from being employed at all, a

consequence of the marriage bar is that many women simply never contemplated getting an education because of the impossible prospect of having both a marriage and a job.

Overtaking discriminatory laws and practices were significant milestones but they did not result in an overnight change in women's participation in science. Despite changes in laws and university admission procedures, it takes time – even generations – for attitudes to change across society (some men still think women's brains are inferior; see Barres 2006), and for young women to realise that a career in science might be possible. Women may have secured more opportunities 'but the world is still, by and large, structured on men's terms.' so surviving without the '...masculine experience of autonomy, mobility and freedom...' (Lake 1999, p. 278) was not straightforward. Despite the odds, some women survived and excelled in various ways as freshwater scientists, and we consider some of those individuals in the next section.

Excellence of women in freshwater ecology – the early pioneers

The science of freshwater ecology is a reasonably new field that garnered momentum after 1900 and has progressed in close association with the study of ecology generally. Consequently, little research into freshwater ecology was carried out by anybody before 1900, although the taxonomy of freshwater organisms was well established and much work focussed on their autecology. There was a notable expansion in freshwater research and publications during the early decades of the 20th century, but women were still rather poorly represented. There are some notable contributions from women during those early years, and they are all the more exceptional when we consider the context in which they studied and worked. During the first years that women were able to receive training as scientists and to pursue their own research interests, women must have faced enormous personal, social and academic challenges. Many of the successful women benefitted from supportive men, either as fathers, husbands, supervisors or financial sponsors, but all would have faced hurdles simply because they were women. Carrying out field work would face the additional challenges of what to wear. Societal expectations in the early decades of the 20th century often demanded that women wear long skirts, long sleeves and hats (Warner and Ewing 2002), of a style that are considered unsuitable for wading in shallow rivers and lakes today and that would probably fail most modern risk assessments. Any woman who managed to make contributions to freshwater ecology during this era was indeed exceptional.

Next, we provide vignettes of four women who were active in freshwater ecology, roughly in the middle 50 years of the 20th century. Looking back helps identify what kinds of contribution to freshwater ecology we consider to be 'significant', not least because the true impact of research is often evident only in retrospect. Also, significance may depend on context and so an awareness of the barriers these women faced (as just discussed) will highlight the magnitude of the challenges they had to overcome in making their contributions. These four women were selected because they did something that was remarkable for women at that time and they made significant research contributions to freshwater ecology. There are, of course, many

women with similar stories and we intend no slight on any individual who is not featured here.

Kathleen E. Carpenter (1891–1970)

Kathleen Carpenter was one of the first women scientists in the UK to study the ecology of animals in rivers and lakes (Duigan 2018). Among the first waves of women to receive a university education in the UK, Carpenter entered University College Wales (Aberystwyth) in 1907, which was one of the more progressive institutions of the time. Carpenter was awarded a B.Sc. in 1910 and, subsequently, earned a M.Sc. and Ph.D. at Aberystwyth. Carpenter is accredited with having authored the first freshwater ecology textbook in English, entitled *Life in Inland Waters: with Especial Reference to Animals* (Carpenter 1928). Although books had been published earlier on particular groups of freshwater organisms by men (Miall 1903; Ward and Whipple 1918) and women also began publishing field guides (e.g. Morgan 1930), Carpenter's book was probably the first freshwater text in English to take a strongly ecological and ecosystem perspective; in her words: 'The standpoint of the work is ecological throughout...' (Carpenter 1928, p. viii). Indeed, Charles Elton's (1927) classic text, *Animal Ecology*, had been published the previous year in the same series of textbooks. Carpenter's own pioneering research documented the environmental impact of metal pollution on the freshwater fauna in Welsh streams and established a direct link between mine waste water and ecological damage (e.g. Carpenter 1924). Unusually, Carpenter was also able to study the ecological recovery of degraded streams when local mines became economically unviable. An early champion of the use of science to inform environmental management, her research findings were used by at least one governmental committee on river pollution (Duigan 2018). After she left Wales, Carpenter continued to research environmental toxicity on fish, but appeared to play a stronger role in lecturing and held positions at several universities in North America, including Radcliffe College (USA), before returning to lecture in the UK and the University of Liverpool.

Margaret Ann Chapman (1937–2009)

Margaret Ann Chapman is recognised worldwide as one of New Zealand's leading limnologists, particularly in the field of zooplankton ecology (Green and Boothroyd 1999). Chapman began her university training at the University of Otago and completed her M.Sc. in 1959. Later, Chapman moved to Scotland to study and received a Ph.D. in 1965 from the University of Glasgow. Subsequently, Chapman took up a lectureship at the University of Auckland and was later appointed to the academic staff at the University of Waikato, New Zealand. Much of her research focused on the taxonomy and ecology of freshwater Crustacea, especially amphipods. As with any country that is reasonably new to science, the task of creating taxonomic descriptions and classifications for a host of new species is not trivial. Perhaps one of her best-known publications is a book on the freshwater Crustacea of New Zealand, which she co-authored with two other women (Chapman *et al.* 1976). Chapman was a founding member of both the Australian Society for Limnology (now the Australian Freshwater Sciences Society) and the New Zealand

Limnological Society (now the New Zealand Freshwater Sciences Society), and also a founding member of the University of Waikato Antarctic Research Programme. Remarkably, Chapman was the first woman – in the world – to lead a scientific expedition to the previously forbidden continent of Antarctica (see above) in 1970–1971. This was a 3-week biological survey of the frozen lakes in the Taylor Dry Valley and she was also one of the first woman scientists to visit the Ross Sea Region of Antarctica. There is even a lake, Lake Chapman, near Granite Harbour in the Ross Dependency named in her honour (latitude: –77.0166667, longitude: 162.3833333).

Rosemary Helen Lowe-McConnell (1921–2014)

Rosemary Lowe-McConnell was one of the pioneers of tropical fish ecology (Bruton 1994; Stiassny and Kaufman 2015). Lowe-McConnell worked in the tropical waters of Africa and South America and contributed significantly to our understanding of the ecology, zoogeography, phenology, evolution and taxonomy of tropical fishes. A thread throughout her research was the need to understand the ecology of fishes, so as to ensure their sustainable utilisation. Educated in the UK, Lowe-McConnell was awarded a B.Sc., M.Sc. and D.Sc. from the University of Liverpool. In 1942, Lowe-McConnell began her research career on the staff of the Freshwater Biological Association, UK, but soon moved to start research on African ichthyology in 1945, and, over the following 8 years or so, she worked in all the East African territories of the UK. Her initial project involved surveying the tilapias and their fisheries in Lake Nyasa (now Lake Malawi), work that laid the foundation for many subsequent studies of Malawian cichlids. Employed as a Research Officer in the British Overseas Research Service, Lowe-McConnell helped found the East African Fisheries Research Organisation and continued research on the biology of the tilapias in East African lakes. Field work was arduous and challenging in these situations, perhaps more so for women, and she often had to work on her own, with only the assistance of local fishermen. Nevertheless, Lowe-McConnell explored novel sampling methods, including night sampling, snorkelling and a 1953 SCUBA dive using rocks tucked in her clothing as weights. This was a productive period and she produced many scientific papers including descriptions of new species (Lowe 1955) and a major work on the differences between the substrate-brooding and mouth-brooding species of tilapias that became the basis for the systematics of tilapiine fishes. Her ecological studies effectively combined information of value to both science and the fisheries of African lakes and set the baseline for later assessments of the impact of fishing and other human pressures on food fish populations.

Lowe-McConnell was forced to resign her job in 1953 because she got married. The UK did not lift this marriage bar on the Foreign Service until two decades later (Fig. 1). Remarkably, Lowe-McConnell managed to stay research active for many more years, although research was typically carried out on a voluntary and expenses-only basis, with occasional supplements from contracts, consultancies, teaching assignments and royalties. Nevertheless, Lowe-McConnell continued to work in Africa and began a fish collection from the Okavango Delta (now in the Natural History Museum, London). Later, Lowe-McConnell

spent several years in South America (British Guiana, now Guyana), carrying out foundational fish surveys in remote and unstudied regions, and also expanding her research to include marine fishes. On a later trip to Brazil, Lowe-McConnell made some of the first studies of the synecology of Amazonian fishes in the Mata Grosso region, which is an area of high endemism among fishes. After her husband retired to the UK, Lowe-McConnell continued research as an Associate of the British Museum (Natural History) in London. This was again a productive period during which she wrote many scientific papers and several influential books (e.g. Lowe-McConnell 1975, 1977, 1987), as well as an active period editing and participating in various international programs. In 1997, Lowe-McConnell was awarded the Linnean Medal (Zoology) by the Linnean Society; at the time she was only the sixth woman to receive the medal, compared with 140 men. As she received her medal, Lowe-McConnell uttered ‘Not bad for someone who hasn’t had a job since 1953!’ (Stiassny and Kaufman 2015, p. 1721).

Ruth Myrtle Patrick (1907–2013)

In the USA, Ruth Patrick was one of the early pioneers of freshwater ecology and phycology, and her impressive contributions are well known around the world. Diatoms were the model organisms for her research throughout her career, but she was renowned for taking a multidisciplinary approach and the context for her investigations included taxonomy and basic and applied ecology. Patrick entered Coker College (a women-only institution) and graduated with a B.Sc. in 1929, before moving to the University of Virginia for postgraduate study (M.Sc. 1931, Ph.D. 1934). There were a few female postgraduate students at the University of Virginia in the early 1930s, but none in the undergraduate school, and many male students and staff were aggressively hostile to these women (Patrick 1997). After postgraduate studies, Patrick moved to the Academy of Natural Sciences in 1933 and worked – unpaid – until 1945, when she became a member of staff. Circa 1945, the State of Pennsylvania donated a large sum of money to the Academy on the condition that it be used to support Patrick’s research into how the species composition of diatoms changed with water quality. The Academy were pleased with the donation but did not want Patrick to run the project because she was ‘...a young woman and all young women waste money.’ (Patrick 1997, p. 7). The Academy acquiesced and allowed Patrick to direct the research only when withdrawal of the money was threatened. Thus, in 1947 Patrick founded the Academy’s Limnology Department and in the following year she directed an unprecedented, large-scale and multidisciplinary field study to test the hypothesis that the biological diversity of a stream (including bacteria, algae, protozoa, rotifers, macroinvertebrates and fish) could be used as a measure of pollution. This work (Patrick 1949) established the importance of diversity as a characteristic of streams and established the strong link between water quality and diversity, which continues to be the foundation of modern environmental assessment. There followed a long and fruitful collaboration with industry on the problems of human impacts on freshwaters, with particular support from the DuPont Co. who were unusually concerned about the effects of their company’s waste on rivers (Patrick 1997), and Patrick served on many state and

federal government advisory panels. In addition to an interest in water quality, Patrick made major contributions to diatom taxonomy (Patrick and Reimer 1966), was involved in palaeolimnological research, and her strong interests in ecology theory led to empirical tests of MacArthur and Wilson's ideas about island biogeography (Patrick 1967). Unsurprisingly, Ruth Patrick received many honours and awards over her life time (Bott and Sweeney 2014). Among the early awards, in 1970 she was the 12th woman in over 100 years to be elected to the National Academy of Sciences, USA.

As the inspiring stories of these four women show, all faced substantial gender-defined barriers in their pursuit of science, for example, they were denied employment, denied remuneration for work performed, denied access to research areas and denied leadership roles. Unquestionably, there were many more obstacles and slights that have not been documented. These were remarkable women and they also made significant contributions to freshwater ecology, i.e. their research was of high quality and had a significant impact. Although women's participation in science generally has improved, new hurdles have also emerged. As long as education and employment practices are defined in terms of the masculine condition (Lake 1999), there will be problems for women, as discussed in the next major part of the paper.

Modern barriers to women in freshwater ecology

In the past 100 years, women have secured large changes that give them new social, political and economic rights, including access to male-dominated workplaces such as universities (Lake 1999). The proportions of women attending university and going on to become academics have increased. Nevertheless, there is robust evidence that barriers to women still exist, particularly for disciplines within science and technology. We will review briefly some statistics compiled by Bell (2009). In the natural and physical sciences in Australia, women are awarded more Bachelor degrees than are men, complete more Honours degrees, and have achieved parity with men in Ph.D. completions and in holding Level A (Senior Tutor) academic positions (Bell 2009). However, at senior levels, the proportions of women in science and technology research positions plummet. Only ~10–15% of academics at the level of Senior Lecturer and above are women. At the most senior ranks (e.g. Professor), only 8% are women, and these figures had not changed substantially since the previous report in 1995 (Bell 2009). Although Bell's report is now almost 10 years old, more modern statistics (e.g. Universities Australia 2017) do not separate science and technology researchers from other disciplines, some of which have less difficulty attracting and retaining female staff. Bell's (2009) figures are comparable with studies elsewhere (e.g. the UK), which show that, even though women steadily increased their numbers within the ranks of academia during the early 20th century, the proportions of women at senior levels (across all disciplines) have not changed since the 1960s (Heward 2005). It would seem that women are either leaving careers in science or failing to get promotion to senior ranks.

In response to statistics such as these, some men continue claiming that women lack the ability to succeed as scientists (see Barres 2006 for examples and a particularly poignant

explanation of how this can be just blatant sexism), but the argument that women lack the mental abilities to do science has been debunked (Barres 2006; Bell 2009). These studies and others (e.g. Evans 2005) present evidence that multiple, entrenched barriers stifle women's full participation in science careers. Some of these barriers relate to the problem that workplaces are still designed around men's life experiences and preferences, which make it difficult for women to participate (Lake 1999). However, the historical section above shows that barriers were repeatedly erected to prevent women from advancing in science careers. It is no longer lawful (in most civilised places) to disqualify women from research careers simply because they are women, but have new and more subtle barriers taken the place of flagrant sexism?

Because the historical arguments centred on women somehow being 'incapable', current methods used to measure 'capability', or the quality or impact of researchers, deserve scrutiny. Attempts at quantitative measures are now common, in part because of the relatively modern obsession with ranking researchers (e.g. Highly Cited Researchers: <https://hcr.clarivate.com/>, accessed November 2018), institutions or countries (e.g. Times Higher World University Ranking: <https://www.timeshighereducation.com/world-university-rankings>, accessed November 2018) according to the (alleged) quality of their research (Lynch 2006; Morrish and Sauntson 2016). A variety of measures is used in ranking schemes, but we will focus on one: the use of citation counts of published works or of individuals as measures of research quality or impact, including the various indices derived from citation counts. We will refer to them all as citation metrics. We focus on citation metrics because their use for evaluating the quality of research and researchers has accelerated in the last decade. In our opinion, there is a danger that, unless the research community starts resisting this practice, citation metrics may become the *only* method by which research and researchers are evaluated.

In the sections to follow, we review the evidence that citation metrics are a reliable measure of research quality or impact. Astonishingly, this evidence is exceptionally poor, and the arguments are false. We then show that the ways citation metrics are used diminish women's contributions to research; they also devalue the types of research contributions epitomised by the four women described above. This material about citation metrics is not connected to specific women because we wish to avoid personalising this part of our critique. However, we present evidence from research published in the discipline of ecology broadly, because it is likely that any discrimination within ecology will also apply to its subdisciplines, such as freshwater ecology. We begin by contrasting how scientists typically define research quality or impact with definitions that appear in the literature promoting citation metrics to measure.

What is research 'quality' and 'impact'?

The quality of scientific research is different from its impact. In our view, quality is an *intrinsic* aspect of research that is assessed across a recognised set of disciplinary standards, for example whether the methods are of high quality, statistical analyses correct, and arguments logical (these standards form basic training in science, as evidenced by texts used to train

B.Sc. students). Collectively, these standards allow scientists to decide whether the conclusions of the research are sound and that is why such standards are applied during peer review for publication. Alternatively, impact captures the *effects* of a publication, researcher or body of work; impact can be assessed against *different* standards. For example, high-impact research may significantly advance the research discipline, or result in new inventions or applications, or solve societal problems. Each of these impacts are valuable contributions but require different forms of evaluation. For example, to evaluate whether research has helped advance the discipline requires a definition of scientific progress. Common yardsticks of scientific progress are new discoveries, insights, and improved capacity for successful prediction (Platt 1964), and researchers or publications can be evaluated retrospectively by examining the extent to which they contributed to that progress (e.g. Real and Brown 1991). Scientists can progress their discipline in multiple ways, such as by generating new ideas or by publishing evidence that overturns current theories. Some scientists contribute a body of work over many years that opens up new frontiers and has collective impact (as epitomised by the pioneering women described above). Evaluating whether research has impact that progresses the discipline is therefore difficult. Evaluation of other sorts of impacts (e.g. impacts of new applications) is likewise not straightforward and would obviously use different criteria.

There are several conclusions we can draw from the above material. First, research may be of high quality but have low impact (however impact is measured). In ecology, for example, field experiments that test general hypotheses are usually preceded by surveys of species abundances or environmental variables, which are required to design experiments successfully. Although papers reporting basic data are essential, they seldom break new ground in the discipline. Second, assessors of research quality or impact need to be individuals that are either contributors to the discipline and have expertise in the relevant research area, or are trained historians specialising in the discipline (e.g. Kingsland 2005). Even if untrained individuals can overcome the challenges of technical language, they lack the expertise needed to place research into a context where its contribution to the discipline can be fully understood and evaluated. In ecology, for example, it takes years of reading the literature, and contributing research to that literature, to gain a comprehensive and authoritative understanding of the discipline. Third, research impact can be assessed only in retrospect because the true impact of research ideas or applications are often not evident for decades, particularly in disciplines such as ecology where numerous tests in different locations and ecosystems are required to evaluate the veracity and generality of hypotheses (Kingsland 2005).

In contrast, the literature that uses or promotes citation metrics to examine research (usually termed bibliometrics or scientometrics; Hood and Wilson 2001) often applies different definitions of ‘research impact’ or ‘research quality’. Some define ‘quality’ or ‘impact’ by using words comparable to the definitions above, such as in the following quotation:

[Research impact] is the ‘impact’ of a publication that is most closely linked to the notion of scientific progress – a paper creating a great impact represents a major contribution to

knowledge *at that time* (although its impact may of course alter with time). Is it possible to obtain any absolute or direct measure of the quality, importance, or impact of a publication? The short answer is ‘No’ [Martin and Irvine 1983, p. 70].

At the other extreme, such views are dismissed out of hand, as evidenced by these quotations from Abramo (2018):

Given that in general the ultimate requirement of a publication is that it provides impact on future scientific advancement, quality needs to refer to impact, and the measure of quality and impact would then be synonymous [p. 592].

... the very essence of scientific activity ... is information processing: the science system consumes, transforms, produces, and exchanges ‘information’. Scientists talk to one another, read each other’s papers, and most importantly, they publish scientific papers... Scientists collect and analyse prior knowledge encoded in verbal forms, add value to it – producing new knowledge, which they nearly always encode in papers made accessible to other scientists, and so contribute to further scientific and technical advancement [p. 593].

These claims are illogical and the views they offer of science and scientists are fictional. Unsurprisingly, these statements are not supported by any cogent arguments about (or even superficial familiarity with) research practices in science, even though ‘Scientometrics is grounded in the quantitative analysis of scientific advances, mainly in the area of the ‘research results’, for which it tries to measure impact, for evaluative purposes.’ (Abramo 2018, p. 592). We have supplied these quotations because they illustrate how some citation metrics literature promulgates bizarre views of science that show little engagement with reality.

Debate on what constitutes scientific progress, research quality and impact is extensive, and a full critique is beyond the scope of this paper (for discussion in the citation metrics literature, see Johnes 1987; Lindsey 1989; Ricker 2017). Rather, we dip into this literature to alert scientists that their concepts of research quality or impact differ considerably from those used by many people advocating citation metrics as a way of measuring science and scientists. Indeed, many scientists may have trouble recognising their discipline’s research practices and values in the sometimes nebulous or outlandish language that is used by non-scientists to describe science. Because ‘impact’ and ‘quality’ are sometimes used interchangeably in the citation metrics literature, we do not distinguish them below in our discussion of citation metrics, even though the discussion above suggests that they are inherently different.

Counting citations as alleged measures of quality or impact

Counting the number of citations of a publication, or of a set of publications from a particular person, or calculating the average per publication in a set, was originally proposed as a measure of quality or impact by a linguist called Eugene Garfield (e.g. Garfield 1970), who went on to advocate such measures for the next 40 years. Because simple citation counts were discovered to have various drawbacks, the *h*-index was proposed as an alternative descriptor for individual people (Hirsch 2005). Since then, suggested elaborations, corrections, adjustments or

alternatives to the *h*-index have consumed most letters of the alphabet (Schreiber 2018), and these citation metrics are now commonly used to evaluate publications, people, journals (through journal impact factors) and institutions (see *Formulae for some citation metrics* in the Supplementary material to this paper).

A fundamental claim of the promoters of citation metrics is that they reliably measure the quality or impact of research. Does evidence support that claim? Garfield's original argument used Nobel Prize winners as a measure of research performance and, hence, quality. The citation counts of Nobel Prize winners are much higher than researchers with comparable numbers of published works in the same field but who have not won a Nobel Prize. The ability of algorithms based on citation counts to identify prize winners was described as 'remarkable' (Garfield and Welljams-Dorof 1992, p. 118). Prizes continue to be used to back claims that citation counts are a reliable measure of performance (for examples, see Bornmann and Daniel 2008) but, unfortunately, this reasoning is flawed by at least two mistakes. First, methods used by Nobel Prize Committees to select winners is secret (as is often the case for committees that select prize winners). The Science Citation Index (SCI) has published citation counts since 1967, and Nobel committees may rely on such information in their deliberations. If so, then the correlation between prize winners and citations counts might be better described as a self-fulfilling prophecy than 'remarkable', but the real problem is that we simply do not know whether citation counts feature within prize committee deliberations.

A second, more substantive problem is that Garfield's argument is based on a classic fallacy of reasoning called 'hasty generalisation'. Hasty generalisation occurs when atypical examples are used to further a claim about an entire group (Toulmin *et al.* 1984). Here, the unspoken assumption is that the citation counts of a tiny percentage of scientists who have won Nobel prizes (~0.01% of scientists within those disciplines for which a Prize exists) can be used to draw conclusions about the performance of *all* scientists. This is a nonsense. To illustrate this fallacy, consider what happens when a linear regression model is fit to a largely amorphous cloud of points with a few, very large outlying values. The outliers can create a statistically significant fit of the model, but a relationship determined entirely by a few extreme values is spurious (Quinn and Keough 2002). Citation counts for a set of publications are always highly skewed, with a few very high (and also very low) numbers, and 95% of publications have counts between these extremes (Fig. 2). A linear regression model fit to such data would be misleading, but this is effectively what people do when they claim that the performance of 0.01% of scientists provides a reliable guide to the performance of the other 99.99%. It is curious that this bogus argument has persisted for decades, but only when the warrants of an argument (i.e. its alleged logic) are spelled out in full, as has been done above, are the insidiously deceptive conclusions of fallacies revealed (Toulmin *et al.* 1984).

Other measures (of an individual or institution) used to warrant citation counts as measures of quality or impact are academic rank, qualifications, success at winning research grants, and departmental or university 'prestige' (Bornmann

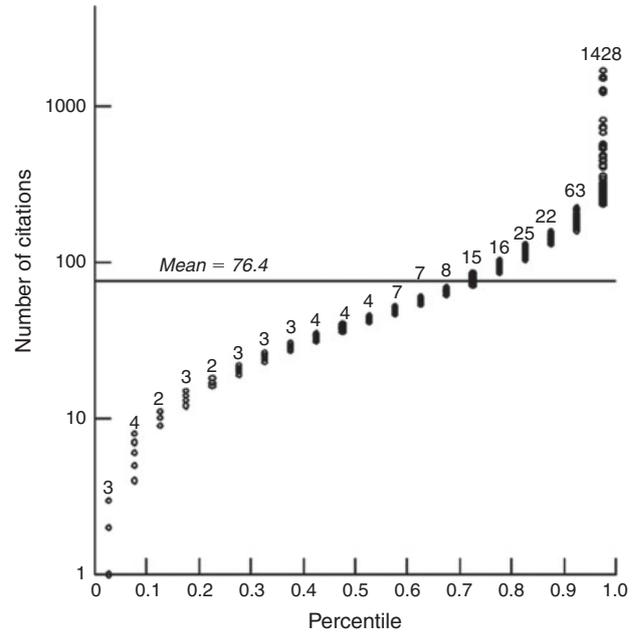


Fig. 2. The number of citations for each of 1383 journal publications plotted on a logarithmic scale against the percentiles for the whole set. Percentiles are in 5% bins. The total range in citation counts was 0–1661, with a mean of 76.4. Numbers over the tops of each set of points within each 5% bin provide the range of citation counts of journal articles within that percentile. The citation counts were downloaded from Web of Science (in October 2018) from journal articles that were searched using 'trophic cascade' (see Table S1, available as Supplementary material to this paper), but without any restrictions on publication year.

and Daniel 2008). These measures can suffer the same problems as described above, but what is worse is that many of these measures are known to have direct, causal relationships with citation counts; this means that they are not independent. For example, citation counts feature in research grant applications, which then influence who is awarded research grants. Thus, claims that these indirect proxies provide *prima facie* evidence that citation counts are a reliable measure of performance commit yet another fallacy, that of circularity (Toulmin *et al.* 1984). Even when it is recognised that alternative, presumed measures of performance, such as prizes, must be strictly independent of citation counts to avoid circularity, strong evidence of that independence is not provided (e.g. Gingras 2014).

Finally, rankings of individuals' research quality gained through peer review are felt to be particularly credible because rankings are (allegedly) independent of citation counts. An early study by Clark (1957) is instructive; he correlated 'eminence rankings' of psychologists with their citation counts to produce a convincing $r = 0.67$. 'Eminence' was initially established using numbers of publications to rank individuals. This list was then sent to 22 eminent people in the field (e.g. heads of professional societies, editors of journals) to add more names of people each judge felt 'in their estimation' should be included (although what strict criteria they used for inclusion is unclear). The longer list was then sent to each person on the list, and they were asked

Table 1. Some factors – other than research quality or originality – that lead to high numbers of citations
Example references come from a 30-year period, showing that these problems have been highlighted over a protracted period

Description	Example references
Paper is an exemplar of poor research or cited because its views are mostly criticised	Bornmann and Daniel (2008)
Publication describes a method (e.g. chemistry procedure, statistical method) that is useful, but in of itself does not lead to insights in the discipline	Lindsey (1989)
Specific research field or topic has higher citation rates than others (e.g. because it has many members) or other time-related factors (e.g. re-discovery of high-quality work)	MacRoberts and MacRoberts (1989); Bornmann and Daniel (2008); Nielsen (2017); Gerow <i>et al.</i> (2018)
Reviews receive more citations than original work	Aksnes (2006); Wendl (2007)
Multiple authors compared to single-authored papers	Leimu and Koricheva (2005)
Journal truncating the number of cited references, meaning that reviews are more likely to be cited than original works	Kostoff (1998); Leimu and Koricheva (2005)
Mistakes in citation and failures to acknowledge sources fully	MacRoberts and MacRoberts (1989); MacRoberts and MacRoberts (2018)
Citations are motivated by social factors (networks, personal acquaintances, flattery to editors or reviewers, persuasiveness to authors' point of view, wishing to be associated with prominent researchers), rather than giving credit for past work	Johnes (1987); MacRoberts and MacRoberts (1989); Kostoff (1998); Leimu and Koricheva (2005); Bornmann & Daniel (2008); MacRoberts and MacRoberts (2018)
Choice of journal, country of publication or residence, institutional 'prestige'	MacRoberts and MacRoberts (1989); Bornmann and Daniel (2008); Leimu and Koricheva (2005); Mirnezami <i>et al.</i> (2016); Bendels <i>et al.</i> (2018)
Characteristics of authors e.g. gender, race, culture, standing in the research community, particularly lead authors	Kelly and Jennions (2006); Symonds <i>et al.</i> (2006); van den Besselaar and Sandström (2016); Bendels <i>et al.</i> (2018)
Self-citations and suggestions that scientists are gaming the citation game (e.g. rotating 1st authorship among groups, 'citation clubs', 'citation exchange')	Johnes (1987); Kostoff (1998); Kelly and Jennions (2006); Macdonald and Kam (2011); Cameron <i>et al.</i> (2016)
Coercion of authors' citations by editors seeking to inflate their journal's impact factor	Wilhite and Fong (2012)

to name top psychologists in their research areas to produce the final list (this ensured all relevant research areas were included). Nepotism clearly creates substantial concerns for the integrity of this process, but a significant problem is that numbers of publications were used to create the first list, which is likely to have had a lasting effect on the rankings as well as having influenced choices of subsequent members. A high correlation between citation counts and numbers of publications for individuals is unsurprising (as we discuss below), and so it reveals little about research quality or impact.

More modern studies using peer review rankings of researchers may reveal only weak correlations with their citation counts ($r = 0.1-0.2$ at best, e.g. Aksnes and Taxt 2004) and these again may be heavily influenced by a few high values (even Clark acknowledged that skewed data were problematic for his correlations). When interviewed about their own papers, scientists do not necessarily regard the number of citations as a guide to the paper's contribution to their discipline (e.g. Aksnes 2006). In the modern era (i.e. following the release of the digital version of the SCI), a large problem is that peer review itself is likely tainted by citation metrics. For example, when outstanding research is defined as '...of great interest with broad impact and with publications in international leading journals...' (Aksnes and Taxt 2004, p. 34), journal impact factors are likely to play a role in such determinations, unless peer review panels are specifically instructed not to use citation metrics either directly or indirectly in their assessments. Even then, given how thoroughly the citation metric disease has infected the

research community (Gingras 2014), it would be difficult for researchers to set aside views that have already been contaminated by citation metrics.

The final nail in the coffin is that citation counts for a publication are strongly influenced by many factors other than quality or impact of the research (Table 1). For example, multi-authored papers receive more citations than single author papers, reviews attract more citations than original work (see next section), and scientists may choose to cite their networks of collaborators (and themselves) in preference to other researchers, even when the latter's research is more appropriate (see review by Bornmann and Daniel 2008). Some modifications to metrics (such as the *h*-index) correct for recognised distortions (review by Waltman 2016), but these modifications tackle problems in a piecemeal fashion. When multiple causal variables (e.g. Table 1) affect a variable (here, citation counts), a multivariate approach is required to isolate the effect created by just one variable (in this case, research quality or impact) from all the background noise (Tabachnick and Fidell 2014). It is otherwise impossible to measure what proportions of citations were caused by the quality or impact of research and it is likely that other factors sometimes have overwhelming effects. Most disturbing of all is growing evidence that authors and editors are deliberately gaming the system. Editors of prominent journals commonly coerce prospective authors into citing papers to inflate the journal's impact factor (Wilhite and Fong 2012), while authors collude in 'citation exchanges' to improve their *h*-index scores (Table 1) and, hence, standing in the research

community. These practices undermine the integrity of academic publishing and make citation metrics little more than a Glass Bead Game^A.

The fatal flaws of using citation counts to measure research quality were highlighted decades ago (Johnes 1987; Lindsey 1989; MacRoberts and MacRoberts 1989) and the criticism continues (see Ricker 2017; MacRoberts and MacRoberts 2018). Nevertheless, these voices have been largely ignored in the stampede to extract information from ‘big data’, so as to analyse alleged research ‘performance’, as we will show later.

How do citation metrics affect women?

We return to the question of whether there are aspects to citation metrics that diminish women’s contributions to research. Citation metrics are clearly an inherently unreliable measure of research quality or impact, but if they are differentially punitive for different members of the research community (e.g. women), then the problems related to their use are far worse. We look at two ways that these differential effects can occur. First, there can be *direct* differences if women suffer discrimination or choose to work differently from men (e.g. a focus on quality rather than quantity); both of these direct effects can translate into differing numbers of publications and, hence, citations. Second, citation metrics may affect women *indirectly* if women and men differ in the proportions of research types they publish that intrinsically attract different rates of citation. We will consider direct effects first because these have had attention in the literature.

Women have lower average values of the *h*-index than do men in science disciplines, including ecology, which on the face of it directly suggests that women’s research attracts fewer citations (Kelly and Jennions 2006; Symonds *et al.* 2006). However, *h*-index values are correlated with total output of publications because researchers with high levels of output tend to be cited more often (this could be caused by lottery or encounter effects in the literature: Kelly and Jennions 2006). Women publish fewer items than men, even when corrections are made for obvious mitigating factors (Leimu and Koricheva 2005; Symonds *et al.* 2006). Differences in publication rates may or may not reflect sexual discrimination, but our concern is not with the causes of this bias but the effects it has on citation metrics. When corrections for numbers of publications are applied to *h*-index values, differences between men and women vanish (Symonds *et al.* 2006). Symonds *et al.* (2006) also found evidence that women publish fewer poorly cited papers, which is consistent with a hypothesis that women may invest more time per article to improve the quality of work rather than focussing on the quantity of publications. In another study of ecologists, Cameron *et al.* (2016) demonstrated that men cite their own publications more frequently than do women (also found by Kelly and Jennions 2006). Elimination of self-citations (as well as correcting for periods of research absence for both men and women) also eliminated gender-biased differences in *h*-index values (Cameron *et al.* 2016). These studies show that

differences between men and women in *h*-index scores reflect different work habits, research priorities or discrimination, and makes unadjusted *h*-index values problematic to interpret.

More insidious problems can occur if citation metrics differentially favour particular types of research and where there are gender-biases in the type of research pursued, which then indirectly lowers numbers of citations. We consider the following two contrasts: empirical *v.* theoretical work and original research *v.* literature reviews.

One obvious difference between empirical and theoretical research projects is how long they can take to complete. Empirical research requires data collection and in ecology this often means fieldwork. Even laboratory work requires collection or maintenance of plants and animals. Many organisms have seasonally affected life cycles, and so data can be collected only at particular times of year, and well-designed field experiments may have to run for years to deliver answers (Underwood 1997). Empirical data may also entail a lot of laboratory time (e.g. processing samples). Alternatively, theoretical research may require mathematics alone or in company with computer-based simulations. Theoretical work is not seasonally restricted, and, unlike humans, computers can make calculations and run simulations continuously until they are completed. Theoretical work should, therefore, often produce results more quickly than most empirical work. Consequently, empiricists probably generate publications more slowly than do theoreticians, with obvious implications for numbers of citations per person. This is a problem in its own right for using citation metrics, but are there gender-based differences in the frequencies of publications in empirical *v.* theoretical research?

In one study, Haller (2014) gathered information from 614 ecologists and evolutionary biologists using an on-line survey to uncover their attitudes to the theoretical-empirical divide in their respective disciplines. He reported the proportions of theoreticians and empiricists that were women, and, although Haller noted that women were more often empiricists than men, he did not test this directly. We used the information in his paper to create the appropriate contingency table (Table 2), which shows that women undertook disproportionately more empirical research than did men (three-quarters of women undertook only empirical work). Women were therefore under-represented in theoretical research in his sample. This is only one study, and it is unclear whether Haller’s sample was representative (because individuals self-nominated to be included in the sample). Nevertheless, the extent of the difference between men and women shows that this matter deserves a lot more investigation. All empiricists are disadvantaged if the time required to complete research is not considered in citation metrics, but, if women generally undertake more empirical research than men, then they will be disadvantaged disproportionately.

Our second comparison is between original research (empirical or theoretical) and reviews that contain no original data. The consensus is that reviews generate more citations (Table 1), but what is the magnitude of the difference? We collected data on citation counts of reviews and articles in three literature

^A*The Glass Bead Game* by Herman Hesse (published in 1943) describes a future in which scholars jockey for position by playing the Glass Bead Game. Rules of the Game are opaque and mysterious. Playing it successfully results in scholarship becoming completely divorced from actual, real-life wisdom or application. Only boys were allowed to play.

Table 2. The frequencies of empirical v. theoretical research conducted by men and women

Numbers of women and men that said they conducted entirely or mostly empirical work, entirely or exclusively theoretical work, or contributed approximately equally to these two types of research ('middle'), based on a study by Haller (2014) of 614 ecology or evolutionary biology researchers in the USA. Four individuals did not reveal their gender or ticked 'other' for this variable, and were grouped with men because Haller did not reveal what types of research they conducted. Statistical analysis of these data shows that women were more likely to conduct empirical (76% of the sample) than theoretical research, or an equal mix of theoretical and empirical work ($\chi^2 = 21.17$, d.f. = 2, two-tailed $P = 0.000025$). The magnitude of this difference is unlikely to be affected by the four individuals added with men

Parameter	Empirical	Middle	Theoretical	Totals
Women	197	37	26	260
Men + four	205	86	63	354
Totals	402	123	89	614

searches, each using terms that capture research on ecological questions of long-standing interest and to which freshwater ecologists have contributed (Table S1, available as Supplementary material to this paper). We tested for an association between publication year and the total numbers of citations for each publication, which allows us to contrast the rate of citation for articles and reviews (Fig. 3, Table 3). In each of the three research areas, the slopes of the lines were the same for reviews and articles, suggesting that both kinds of publications attract citations at similar rates. However, reviews gathered approximately three times more citations than did articles in each research area (Table 3). If this difference is general to other areas of ecology (and science at large), then researchers who publish few or no reviews are significantly disadvantaged in citation counts. We think that this acts as a strong disincentive to researchers to gather empirical data that require lengthy effort. Declines in production of such empirical work (and increases in quicker and cheaper forms of research such as data mining and modelling) have already been documented (Lindenmayer and Likens 2011), as have declines in recognition that truly innovative empirical work often takes >5 years to produce (Statzner and Resh 2010). The differential numbers of citations flowing to reviews is simply another incentive to avoid lengthy empirical work. An obvious solution is that the citation counts of reviews should be divided by three so as to achieve parity with original research. We call this the downward levelling (DL) correction because it levels the playing field.

Are there differences between men and women in whether they gather citations from reviews? Studies have revealed differences in frequencies of lead and senior authorship (and also subsequent effects on citations) between men and women in science (e.g. Bendels *et al.* 2018), but we are unaware of any specific research on ecologists that addresses this question. We collected data from our literature searches (see Table S1) to test whether there were gender-biases in authorship of reviews v. original articles. The results showed that men were lead authors more often than women overall (76%); however, men lead reviews at a significantly higher frequency (86%) than they lead articles (67%; Table 4). Overall, the proportions of reviews

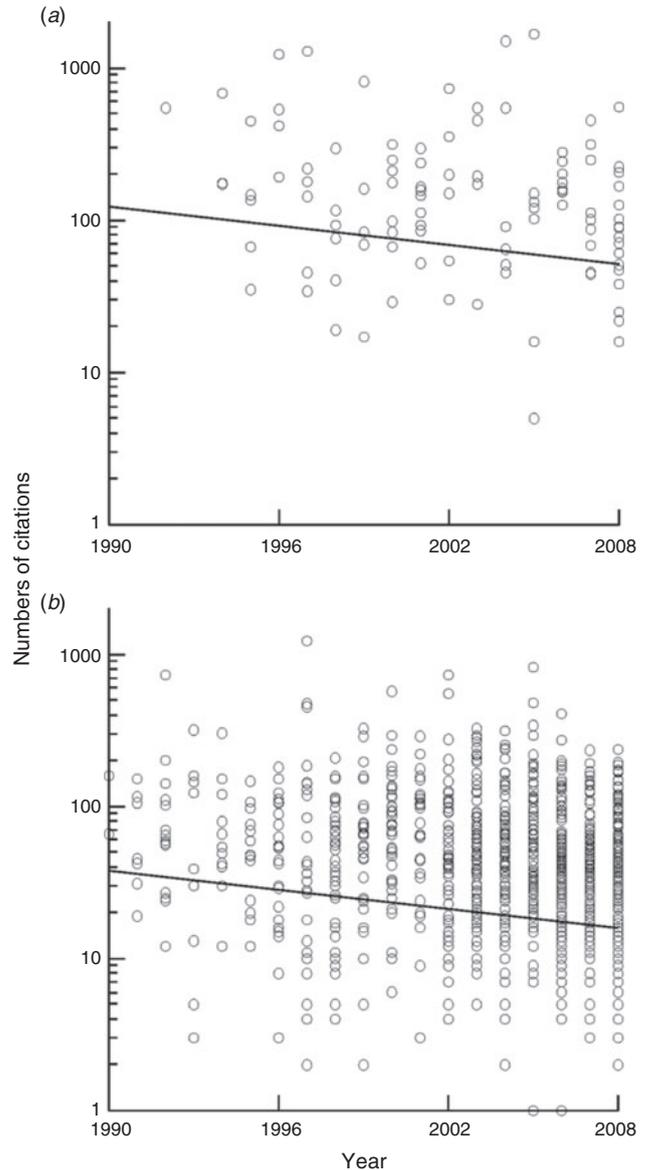


Fig. 3. The number of citations of journal publications that are either (a) reviews or (b) original articles v. the year of publication for publications on 'trophic cascades'. Journal publications were gathered as per Table S1, available as Supplementary material to this paper. The lines are least-square linear regressions performed on \log_{10} -transformed data.

and articles with none v. at least one female author were not significantly different ($\chi^2 = 1.58$, $P = 0.21$) however the total numbers of female authors differed between publication types. Women were significantly less often authors of reviews (80% of which were multi-authored) than of articles (Table 4). When publications were multi-authored and with at least one woman and one man, reviews again more often were led by men (70%), but women were more often lead authors on articles (60%; Table 4). These results are preliminary, given the search was undertaken only on two, albeit general, topics, and sample sizes are relatively modest, but our findings are consistent with those of other studies (e.g. Bendels *et al.* 2018). They suggest that men

Table 3. A comparison of numbers of citations gathered by reviews v. articles

Outcomes of regression analyses on \log_{10} -transformed numbers of citations v. the year of publication for reviews and articles for three sets of journal publications, using three different search terms. In all cases, the slopes of regression lines were the same for articles and reviews, meaning that the y -intercepts can be directly compared to examine differences in average numbers of citations (for details, see Table S1, available as Supplementary material to this paper). Provided in the table are the search terms, the value of the y -intercept, the y -intercept back-calculated onto an ordinary numerical scale (rounded to nearest whole digit) and the ratio of average number of citations of articles to that of reviews

Search term	Slope of line	y -intercept (log scale)		y -intercept (numerical scale)		Ratio
		Reviews	Articles	Reviews	Articles	
Trophic cascade	-0.021	2.070	1.582	117	38	3.08
Density-depend* OR density depend*	-0.015	1.958	1.485	91	29	3.14
Disturbance AND species diversity	-0.021	1.995	1.458	99	29	3.41

Table 4. A comparison of frequencies of authorship by men and women of reviews v. articles

Numbers of journal publications according to type of paper (review v. article) compared to numbers of publications categorised into three, separate contingency tables according to: whether lead authors were men v. women; the total numbers of men and women authors; and for mixed research teams (defined as having at least one member of each gender), numbers of publications where lead authors were men v. women. For each contingency table, Chi-Square (with one degree of freedom) was used to test whether the characteristic of authors differed between reviews and articles. See Table S1 (available as Supplementary material to this paper) for details of literature searches and methods. Asterisks indicate significance with α of 0.05

Publication type	Lead author male	Lead author female	Number of men	Number of women	Totals	Test outcome
Gender of lead author						
Articles	44	23			67	$\chi^2 = 7.52, P = 0.006^*$
Reviews	60	10			70	
Col. totals	104	33			137	
Total numbers of men and women authors						
Articles			142	49	191	$\chi^2 = 4.02, P = 0.045^*$
Reviews			179	38	217	
Col. totals			321	87	408	
Gender of lead author of mixed research teams						
Articles	10	16			26	$\chi^2 = 4.74, P = 0.029^*$
Reviews	16	7			23	
Col. totals	26	23			49	

are more likely to initiate reviews or be invited to write them, and, when reviews are multi-authored, men collaborate with male colleagues more frequently than they do when working on articles. Again, our interest is not in what causes the discrepancy, but the implications it has for citation metrics. If women are authors or co-authors of significantly fewer reviews, then they will gather fewer citations than do men. There could be other gender-biased differences in types of research that also affect counts of citations; this topic needs a lot more research.

Citation metrics are flawed and are damaging research

We have demonstrated that citation counts do not provide reliable measures of the quality or impact of research. Even allowing for different definitions of 'quality' and 'impact' in the citation metrics literature (and the sometimes hopeless muddling of these different aspects), the evidence that citation metrics measure either of these attributes is fatally flawed by illogical reasoning. Additionally, although scientists certainly cite research of good quality to back their arguments, they also cite research for many other reasons that have nothing to do with research quality. Most damning of all, citation metrics undermine women – and also young researchers and people from non-English speaking

backgrounds (Table 1). The obvious conclusion is that these citation metrics should be abandoned as a reliable measure of quality or impact of research or researchers.

The citation metrics literature does not see it that way. A simple search in Web of Science (conducted in October 2018) using 'citation count' or 'h-index' or 'journal impact factor' as search terms produced >3700 papers published since 1972. An astonishing 90% of these papers have been published in just the past 10 years. Only a third of papers were published in medical, science or social science journals (i.e. written by people who are likely to have both knowledge of, and training in, the relevant discipline) and these papers were spread across many different medical, science and social science research topics (so there were only a few papers on any one area). The other two-thirds of papers were published in journals classified by Web of Science as Information Science/Library Science, Computer Science or Information Systems. Certainly, some of these publications criticise citation metrics but they are a minority. Many papers use citation metrics to evaluate journals, researchers or institutions in a chosen research discipline, often with the specific goal of ranking those entities. Such publications can be found even in Library Science journals (e.g. Bapte and Gedam 2018; Bhui and

Sahu 2018; Nanda *et al.* 2018; Shao *et al.* 2018; Elango 2019), which many researchers might otherwise assume would eschew such work.

This practice of using citation metrics to ‘evaluate’ research and researchers has flourished because the fatal flaws with citation metrics are being deliberately ignored. MacRoberts and MacRoberts (2018), who had criticised citation metrics 30 years earlier (MacRoberts and MacRoberts 1989), suggested that it is a combination of problems. In short, many people promoting citation metrics lack relevant scientific training, appear not to understand methods for data collection and sampling, and they make elementary mistakes in reasoning that they do not even recognise. Likewise, they fail to grasp that low numbers of citations do not indicate that publications are ‘unimportant’, they assume that all citations are of equal value in signalling ‘quality’, and they do not have the expertise to read the publications to assess whether or not this assumption is true. Moreover, although they may acknowledge ‘limitations’ to citation metrics, these limitations are never permitted to stand in the way of their continued use. A lack of evidence that citations are a reliable measure of quality or impact is constantly either dodged (by citing studies that themselves simply claimed this was true but without evidence) or simply waved away. As MacRoberts and MacRoberts (2018) put it:

What we witness here is how blind ideologues can be; they simply dismiss—or ignore—data contradicting their beliefs, and their theories and opinions increasingly take precedence over the facts [p. 479].

Distressingly, citation metrics are nevertheless used to decide which research fields within a particular discipline should be supported (Morrish and Sauntson 2016) and are advocated for use by human resource managers to get more out of their academic staff (Jaskiene 2015). Some papers even present computer programs that use citation counts and machine learning to rank researchers and suggest that these programs can replace independent peer review to decide who gets research funding (Ebadi and Schiffauerova 2016). In the past 10 years, the *h*-index and some of its variations have been normalised as measures of research and researcher quality. These faulty, deceptive numbers are used by some institutions to measure performance of researchers and create ‘league tables’ in ways that compromise academic freedom, de-value some types of research, and create great stress, even leading to suicide (Burrows 2012; Morrish and Sauntson 2016). As one pair of authors put it:

This is acanemia, where etiolated, dressage trained academics ... shuffle round meeting their targets, brandishing their *h*-indices, but joyless and insecure [Morrish and Sauntson 2016, p. 61].

If citation metrics are permitted to take over as the main indicators of ‘quality or impact’ (as some bibliometricians, scientometricians and research administrators are pushing for), the only ‘dressage trained academics’ in acanaemia will be the show ponies who are successful at playing the Glass Bead Game.

It is not hyperbole to suggest that the unfettered application of citation metrics has the potential to do great damage to research and researchers. Corrections to existing citation metrics exist but they are multitudinous and there is no

agreement about which ones are best (Schreiber 2018). Commonly used metrics do not address the basic problem that some essential research will generate publications slowly or gather few citations nor do they recognise that researchers are willing to behave unethically to game their scores (Table 1). The same criticisms apply to journal impact factors (JIF). In ecology, journals dedicated to publishing excellent original research, especially empirical research, are ranked lower than journals that publish mostly reviews or opinion pieces. Such rankings suggest that empirical research is ‘less important’, which is patently absurd. Moreover, highly skewed distributions of citations (Fig. 2) mean that JIF scores are largely determined by an unrepresentative and tiny number of very highly cited papers (Schreiber 2018), thus producing misleading values. When we consider also that editors of ‘highly ranked’ journals are deliberately coercing authors into inflating their journals’ JIF scores (Table 1), it renders a view that journal impact factors verge on meaningless. We find this situation reprehensible and unacceptable.

Research is a multi-dimensional activity with many qualitative and essential differences among researchers. An analogy is to consider that building a piece of fine furniture requires a diverse set of tools (e.g. chisels, saws, drills), materials (e.g. wood, glue, hinges, wax), and diverse skills (preparing timber, making dove joints, applying French polish). The quality of furniture cannot be measured solely by the size of the screwdriver. Freshwater ecology, and other science disciplines, are much the same in that they can progress only with contributions from diverse kinds of information (field experiments or surveys, laboratory research, theory, computer simulations, meta-analyses), contributors with diverse skill sets (e.g. taxonomy, experimental design, statistical analyses, field ecology, modelling, biomathematics), and different perspectives and interests (e.g. systematics, basic ecology, applied ecology, hydrology and hydraulics). Scientific progress requires all the parts, but mindless use of citation metrics will damage this, perhaps fatally. Taxonomists may be the first to disappear (their numbers have dwindled already) because taxonomic papers typically get very few citations (for a good reason; MacRoberts and MacRoberts 2018) but ecology cannot advance without taxonomy and systematics. This is particularly so for freshwater ecology, where many invertebrate species have never been collected or described. Likewise, meta-analyses in ecology can test the generality of models (across different places or ecosystems) but are impossible without a wealth of empirical research to examine. As described above, empirical research takes time to produce and also generates fewer citations than do reviews. Full-time empiricists and many other types of researchers, no matter how talented, will not survive in acanaemia.

All of this is bad enough, but the citation metrics axe will clearly fall disproportionately upon women. As reviewed above, the *h*-index is biased against women directly and also indirectly because women undertake disproportionately more research that generates fewer citations simply because of its nature. The answer to this problem does not lie in pushing women to give up empirical work and switch to writing more reviews (advice that the authors of this paper have been given by men). Choice of research approach should be dictated by the interests and talents of the researcher and the needs of the discipline, not

simply to boost the citation counts of authors - otherwise, we are reduced to playing the Glass Bead Game instead of doing science.

The future

We began this paper by discussing a topic that seems remote from citation metrics – the ridiculous reasons used by men in the 19th and early 20th centuries to bar women from education and research careers. In the 21st century, women enjoy equal rights to employment and education under the law, in many places. It is tempting to consign the blatant sexism of the past to history and to consider all that irrelevant now. However, to do so is to forget that men's life experiences still largely structure workplace organisation and values (Lake 1999), and universities are prime examples (Evans 2005). We do not suggest that anybody created citation metrics *deliberately* to discriminate specifically against women. Rather, citation metrics are an outcome of assuming that high numbers of citations are a natural expression of 'quality' coming to the fore. MacRoberts and MacRoberts (2018) likened the approach of evaluating researchers using citation metrics to the way baseball players are ranked using batting averages^B. It is the same type of simplistic thinking; it reduces complex achievements down to a single number that enables data-crunchers and bean-counters – who then claim the mantle of 'expert' even when they have no expertise in the research field – to identify supposed research 'elites'. Citation exchanges and other unethical practices that deliberately game the system to reinforce the position of these so-called elites are ignored (Macdonald and Kam 2011). The skewed few, as MacDonald and Kam termed them, are then happy to support the view that their high numbers of citations distinguish them from the rabble. We see an analogy with the arguments raised by wealthy, upper-class men to prevent women (and also working-class men and women) from attending university. Those arguments ensured that only people of the 'right' station and gender were entitled to an education, thus securing the position of the ruling classes (Hubbard 1990). Viewed from this perspective, sexual discrimination caused by the unthinking use – and acceptance – of citation metrics looks like a modern version of an old problem. Of course, women are not banned from research any more, but insufficient numbers of women hold senior research positions, which means that women have had little voice in the debate about how to (or even whether to) measure the quality or impact of research using citations. As such, we think it is unsurprising that citation metrics appear to reflect the probable biases, life experiences and preferences of men.

The language in the above paragraph is blunt, but our wish is to jolt researchers out of complacency about citation metrics. We think it is critical that more of the research community recognises the terrible web of mismeasurement that is ensnaring all of us. How can we defeat the scourge of poorly conceived citation metrics? The ideal outcome would be for citation metrics to be abandoned altogether, but that seems unlikely. Nevertheless, there are constructive ways forward. First, researchers should demand that citation metrics must

demonstrably measure quality or impact according to definitions of those words with which the research community agrees. As explained earlier, these terms have been muddled, and some definitions of impact used in the citation metrics literature have nothing to do with some research outcomes. Second, editors and authors that engage in unethical practices must be exposed by citation metrics that are robust to attempts to game the system. Third, citation metrics must not discriminate against women (or any other group), and those promoting these measures should be *required* to demonstrate *a priori* that they are not discriminatory. It should not be left to researchers, such as us, to point out these damaging mistakes retrospectively. Non-discriminatory corrections have been suggested (e.g. Symonds *et al.* 2006) but these need to be coupled with corrections for different types of research. Nobody should be penalised because they choose to undertake essential research that takes years to complete and generates citations only slowly. The DL correction (above) revises citation counts so that original work is not penalised compared with reviews, and we suggest this is a good start. Nevertheless, far more investigation is required into inequities in citation counts created by the nature of research, especially where there are gender (or other) biases in who undertakes the work. Ultimately, citation metrics must compare like with like and not bundle together qualitatively different types of research that guarantees some types of researchers will always be ranked at the bottom. Finally, we think the entire research community and especially those at senior levels of leadership (Vice-Chancellors, Deans, and Heads of Schools) need to communicate to bureaucrats, administrators and managers that commonly used citation metrics are deeply flawed and are in urgent need of revision. Even then, citation metrics cannot replace a cogently argued case about how an applicant's research has contributed to their field, nor can it replace the insights of peer review.

Finally, we can look to the past for inspiration. In ~140 years, women have gone from first stepping through the university door to having fully fledged careers in all branches of science and technology. The four women pioneers we described above were great not because they had many publications or because their publications were cited many times, although this may be true. They were great because they defied conventions and broke barriers, thus paving the way for women (and men!) who followed and made really important contributions to freshwater ecology and ecology more broadly. Their stories enable us to re-focus attention away from citations and back to what it means to be a great scientist. One message for the next generation is this: science is not a popularity contest. Contributions cannot be measured solely by citations (and absolutely not by re-tweets!). Ground-breaking scientists must have the passion and commitment to pursue ideas in the face of opposition and disapproval, as did our four inspiring women. The quality of that science must not be compromised by cutting corners or refusing to gather data needed to enable research that will deliver definitive answers to general questions. Our four inspiring women produced basic building blocks of taxonomic species descriptions and autecologies (Chapman, Lowe-McConnell, Patrick), carried out ground-breaking empirical field work,

^BBatting average in baseball is a measure of batting ability and is the number of hits where the batter successfully reaches a base divided by the total number of times at bat.

often in remote locations (all four), and developed new methods (Lowe-McConnell, Patrick). This basic research is what enabled them to contribute fresh solutions to important applied problems of anthropogenic impacts (Carpenter, Patrick) and sustainable fisheries (Lowe-McConnell), test hypotheses in basic ecology and evolution (Lowe-McConnell, Patrick), synthesise vast amounts of material into textbooks suitable for teaching or learned monographs that we still rely on today (all four) ... and we could go on. An important point is that in ecology (and perhaps other fields), significant contributions often come from whole bodies of work like these, not individual publications. Many of these significant contributions would not be detected – in fact they would be discouraged and diminished – by citation metrics, and yet these kinds of contributions are needed now just as much as they were needed then. All these considerations mean that young scientists should look for mentors among their research peers and avoid taking advice about their careers from people outside their research field, and, most especially, if they are peddling citation metrics. In closing, we think it is important to remember and live by a pair of old sayings. First, none of us can take absolute credit for our discoveries, because we all stand on the shoulders of those who went before. Second, those people unfamiliar with the history of their discipline are doomed to make and suffer from the mistakes of the past.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Declaration of funding

This research did not receive any specific funding.

Acknowledgements

The authors thank Rebecca Lester for the invitation to contribute a paper to this special issue of Marine and Freshwater Research. B. J. Downes thanks the Faculty of Science and School of Geography at the University of Melbourne for granting a sabbatical, which has made the timely production of this paper possible. We are very grateful to our two referees whose careful, thoughtful comments helped us improve the manuscript.

References

- Abramo, G. (2018). Revisiting the scientometric conceptualization of impact and its measurement. *Journal of Informetrics* **12**, 590–597. doi:10.1016/J.JOI.2018.05.001
- Aksnes, D. W. (2006). Citation rates and perceptions of scientific contribution. *Journal of the American Society for Information Science and Technology* **57**, 169–185. doi:10.1002/ASI.20262
- Aksnes, D. W., and Taxt, R. E. (2004). Peer reviews and bibliometric indicators: a comparative study at a Norwegian university. *Research Evaluation* **13**, 33–41. doi:10.3152/147154404781776563
- Bapte, V. D., and Gedam, J. (2018). A scientometric profile of Sant Gadge Baba Amravati University, Amravati during 1996–2017. *DESIDOC Journal of Library and Information Technology* **38**, 326–333. doi:10.14429/DJLIT.38.5.13194
- Barres, B. A. (2006). Does gender matter? *Science* **442**(7099), 133–136.
- Bell, S. (2009). 'Women in Science: Maximising Productivity, Diversity and Innovation.' (Federation of Australian Scientific and Technological Societies: Canberra, ACT, Australia.)
- Bendels, M. H. K., Müller, R., Brueggmann, D., and Groneberg, D. A. (2018). Gender disparities in high-quality research revealed by Nature Index journals. *PLoS One* **13**(1), e0189136. doi:10.1371/JOURNAL.PONE.0189136
- Bhui, T., and Sahu, N. B. (2018). Publications by faculty members of humanities and social science departments of IIT Kharagpur: a bibliometric study. *DESIDOC Journal of Library and Information Technology* **38**, 403–409. doi:10.14429/DJLIT.38.6.13569
- Bornmann, L., and Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior. *The Journal of Documentation* **64**, 45–80. doi:10.1108/00220410810844150
- Bott, T. L., and Sweeney, B. W. (2014). 'Ruth Patrick 1907–2013. A Biographical Memoir.' (National Academy of Science: Washington, DC, USA.)
- Bruton, M. N. (1994). The life and work of Rosemary Lowe-McConnell: pioneer in tropical fish ecology. *Environmental Biology of Fishes* **41**, 67–80. doi:10.1007/BF02197835
- Burrows, R. (2012). Living with the h-index? Metric assemblages in the contemporary academy. *The Sociological Review* **60**(2), 355–372. doi:10.1111/J.1467-954X.2012.02077.X
- Cameron, E. Z., White, A. M., and Gray, M. E. (2016). Solving the productivity and impact puzzle: do men outperform women, or are metrics biased? *Bioscience* **66**, 245–252. doi:10.1093/BIOSCI/BIV173
- Carpenter, K. E. (1924). A study of the fauna of rivers polluted by lead mining in the Aberystwyth district of Cardiganshire. *Annals of Applied Biology* **11**, 1–23. doi:10.1111/J.1744-7348.1924.TB05689.X
- Carpenter, K. E. (1928). 'Life in Inland Waters: with Especial Reference to Animals.' (Sidgwick & Jackson: London, UK.)
- Chapman, A., Lewis, M. H., and Stout, V. M. (1976). 'Introduction to the Freshwater Crustacea of New Zealand.' (Collins: Auckland, New Zealand.)
- Chipman, E. (1986). 'Women on the Ice: a History of Women in the Far South.' (Melbourne University Press: Melbourne, Vic., Australia.)
- Clark, K. E. (1957). 'America's Psychologists: a Survey of a Growing Profession.' (American Psychological Association: Washington, DC, USA.)
- Clarke, E. H. (1874). 'Sex in Education: or, a Fair Chance for the Girls.' (James R Osgood and Co.: Boston, MA, USA.)
- Duigan, C. (2018). Who was ... Kathleen Carpenter? *The Biologist* **65**(3), 22–25.
- Ebadi, A., and Schiffauerova, A. (2016). iSEER: an intelligent automatic computer system for scientific evaluation of researchers. *Scientometrics* **107**, 477–498. doi:10.1007/S11192-016-1852-2
- Elango, B. (2019). A bibliometric analysis of literature on engineering research among BRIC countries. *Collection and Curation* **38**, 9–14. doi:10.1108/CC-05-2017-0017
- Elton, C. S. (1927). 'Animal Ecology.' (Sidgwick & Jackson: London, UK.)
- Evans, M. (2005). 'Killing Thinking: the Death of the Universities.' (Bloomsbury Publishing: London, UK.)
- Garfield, E. (1970). Citation indexing for studying science. *Nature* **227**, 669–671. doi:10.1038/227669A0
- Garfield, E., and Welljams-Dorof, A. (1992). Of Nobel class: a citation perspective on high impact research authors. *Theoretical Medicine* **13**, 117–135. doi:10.1007/BF02163625
- Gerow, A., Hu, Y., Boyd-Graber, J., Blei, D. M., and Evans, J. A. (2018). Measuring discursive influence across scholarship. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 3308–3313. doi:10.1073/PNAS.1719792115
- Gingras, Y. (2014). Criteria for evaluating indicators. In 'Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact'. (Eds B. Cronin and C. R. Sugimoto.) pp. 109–125. (MIT Press: Cambridge, MA, USA.)
- Green, J., and Boothroyd, I. (1999). Ann Chapman: inspirational limnologist. *New Zealand Journal of Marine and Freshwater Research* **33**, 333–340. doi:10.1080/00288330.1999.9516880

- Haller, B. (2014). Theoretical and empirical perspectives in ecology and evolution: a survey. *Bioscience* **64**, 907–916. doi:10.1093/BIOSCI/BIU131
- Heward, C. (2005). Women and careers in higher education: what is the problem? In 'Breaking Boundaries: Women in Higher Education'. (Eds L. Morely and V. Walsh.) pp. 9–22. (Taylor & Francis: London, UK.)
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 16569–16572. doi:10.1073/PNAS.0507655102
- Hood, W. W., and Wilson, C. S. (2001). The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics* **52**, 291–314. doi:10.1023/A:1017919924342
- Hubbard, R. (1990). 'The Politics of Women's Biology.' (Rutgers University Press: New Brunswick, NJ, USA.)
- Jaskiene, J. (2015). HRM practices enhancing research performance. *Procedia: Social and Behavioral Sciences* **213**, 775–780. doi:10.1016/J.SBSPRO.2015.11.474
- Johnes, G. (1987). Research performance indications in the university sector. *Higher Education Quarterly* **42**, 54–71. doi:10.1111/J.1468-2273.1987.TB02114.X
- Jones, B. (2012). Women won't like working in Antarctica as there are no shops and hairdressers. In *The Telegraph*, 20 May 2012. Available at www.telegraph.co.uk/news/earth/environment/9260864/Women-wont-like-working-in-Antarctica-as-there-are-no-shops-and-hairdressers.html [Verified 4 November 2018].
- Kelly, C. D., and Jennions, M. D. (2006). The *h* index and career assessment by numbers. *Trends in Ecology & Evolution* **21**, 167–170. doi:10.1016/J.TREE.2006.01.005
- Kingsland, S. E. (2005). 'The Evolution of American Ecology, 1890–2000.' (John Hopkins University Press: Baltimore, MD, USA.)
- Kostoff, R. N. (1998). The use and misuse of citation analysis in research evaluation. *Scientometrics* **43**, 27–43. doi:10.1007/BF02458392
- Lake, M. (1999). 'Getting Equal: the History of Australian Feminism.' (Allen & Unwin: Sydney, NSW, Australia.)
- Leimu, R., and Koricheva, J. (2005). What determines the citation frequency of ecological papers? *Trends in Ecology & Evolution* **20**, 28–32. doi:10.1016/J.TREE.2004.10.010
- Lindenmayer, D. B., and Likens, G. E. (2011). Losing the culture of ecology. *Bulletin of the Ecological Society of America* **92**, 245–246. doi:10.1890/0012-9623-92.3.245
- Lindsey, D. (1989). Using citation counts as a measure of quality in science. Measuring what's measurable rather than what's valid. *Scientometrics* **15**(3–4), 189–203. doi:10.1007/BF02017198
- Lowe, R. H. (1955). New species of *Tilapia* (Pisces, Cichlidae) from Lake Jipe and Pangani River, East Africa. *Bulletin of the British Museum (Natural History). Historical Series* **2**, 349–368.
- Lowe-McConnell, R. H. (1975). 'Fish Communities in Tropical Freshwaters: their Distribution, Ecology, and Evolution.' (Longman: London, UK.)
- Lowe-McConnell, R. H. (1977). 'Ecology of Fishes in Tropical Waters.' (Edward Arnold: London, UK.)
- Lowe-McConnell, R. H. (1987). 'Ecological Studies in Tropical Fish Communities.' (Cambridge University Press: New York, NY, USA.)
- Lynch, K. (2006). Neo-liberalism and marketisation: the implications for higher education. *European Educational Research Journal* **5**, 1–17. doi:10.2304/EERJ.2006.5.1.1
- Macdonald, S., and Kam, J. (2011). The skewed few: people and papers of quality in management studies. *Organization* **18**, 467–475. doi:10.1177/1350508411403533
- MacRoberts, M. H., and MacRoberts, B. R. (1989). Problems of citation analysis: a critical review. *Journal of the American Society for Information Science* **40**, 342–349. doi:10.1002/(SICI)1097-4571(198909)40:5<342::AID-AS17>3.0.CO;2-U
- MacRoberts, M. H., and MacRoberts, B. R. (2018). The mismeasure of science: citation analysis. *Journal of the Association for Information Science and Technology* **69**, 474–482. doi:10.1002/ASI.23970
- Martin, B. R., and Irvine, J. (1983). Assessing basic research: some partial indicators of scientific progress in radio astronomy. *Research Policy* **12**, 61–90. doi:10.1016/0048-7333(83)90005-7
- Mason, J. (1992). The admission of the first women to the Royal Society of London. *Notes and Records of the Royal Society of London* **46**, 279–300. doi:10.1098/RSNR.1992.0027
- Miall, L. C. (1903). 'The Natural History of Aquatic Insects.' (Macmillan and Co.: London, UK.)
- Mirnezami, S. R., Beaudry, C., and Larivière, V. (2016). What determines researchers' scientific impact? A case study of Quebec researchers. *Science & Public Policy* **43**, 262–274. doi:10.1093/SCIPOL/SCV038
- Morgan, A. H. (1930). 'Field Book of Ponds and Streams.' (G. P. Putnam's Sons: New York, NY, USA.)
- Morrish, L., and Sauntson, H. (2016). Performance management and the stifling of academic freedom and knowledge production. *Journal of Historical Sociology* **29**, 42–64. doi:10.1111/JOHS.12122
- Nanda, S., Mishra, M., and Ramesh, D. B. (2018). Performance analysis and ranking of corporate medical institutions in India. *DESIDOC Journal of Library and Information Technology* **38**, 342–348. doi:10.14429/DJLIT.38.5.13000
- Nielsen, M. W. (2017). Scientific performance assessments through a gender lens: a case study on evaluation and selection practices in academia. *Science & Technology Studies* **31**, 2–30. doi:10.23987/STS.60610
- Patrick, R. (1949). A proposed biological measure of stream conditions, based on a survey of the Conestoga Basin, Lancaster County, Pennsylvania. *Proceedings of the Academy of Natural Sciences of Philadelphia* **101**, 277–341.
- Patrick, R. (1967). The effect of invasion rate, species pool, and size of area on the structure of the diatom community. *Proceedings of the National Academy of Sciences of the United States of America* **58**, 1335–1342. doi:10.1073/PNAS.58.4.1335
- Patrick, R. (1997). The development of the science of aquatic ecosystems. *Annual Review of Energy and the Environment* **22**, 1–11. doi:10.1146/ANNUREV.ENERGY.22.1.1
- Patrick, R., and Reimer, C. W. (1966). The diatoms of the United States, exclusive of Alaska and Hawaii. Number 13. In 'Monographs of the Academy of Natural Sciences of Philadelphia'. (Academy of Natural Sciences of Philadelphia: Philadelphia, PA, USA.)
- Pennington, C. (2015). 'The Historic Role of Women Scientists at BGS and a Look at What Is Happening Today.' Open Research Archive 514086. (Natural Environment Research Council: Swindon, UK.)
- Platt, J. R. (1964). Strong inference. *Science* **146**, 347–353. doi:10.1126/SCIENCE.146.3642.347
- Quinn, G. P., and Keough, M. J. (2002). 'Experimental Design and Data Analysis for Biologists.' (Cambridge University Press: Cambridge, UK.)
- Real, L. A., and Brown, J. H. (Eds) (1991). 'Foundations of Ecology: Classic Papers with Commentaries.' (University of Chicago Press: Chicago, IL, USA.)
- Ricker, M. (2017). Letter to the editor: about the quality and impact of scientific articles. *Scientometrics* **111**, 1851–1855. doi:10.1007/S11192-017-2374-2
- Sawer, M. (1996). 'Removal of the Commonwealth Marriage Bar: a Documentary History'. (Centre for Research in Public Sector Management, University of Canberra: Canberra, ACT, Australia.)
- Schreiber, M. (2018). A skeptical view on the Hirsch index and its predictive power. *Physica Scripta* **93**, 102501. doi:10.1088/1402-4896/AAD959
- Shao, Z. Y., Li, Y. M., Ke, W., Guo, Y. J., Fan, F., Fen, H., Nui, Y. F., and Yang, Z. (2018). How academic librarians involve and contribute in

- research activities of universities? A systematic demonstration in practice through comparative studies of research productivities and research impacts. *Journal of Academic Librarianship* **44**, 805–815. doi:10.1016/J.ACALIB.2018.09.015
- Statzner, B., and Resh, V. H. (2010). Negative changes in the scientific publication process in ecology: potential causes and consequences. *Freshwater Biology* **55**, 2639–2653. doi:10.1111/J.1365-2427.2010.02484.X
- Stiassny, M. L. J., and Kaufman, L. S. (2015). Rosemary Lowe-McConnell, obituary. *Environmental Biology of Fishes* **98**, 1719–1722. doi:10.1007/S10641-015-0409-1
- Sudgen, D. (1987). The polar and glacial world. In 'Horizons in Physical Geography'. (Eds M. J. Clark, K. J. Gregory, and A. M. Gurnell.) pp. 214–231. (Macmillan Education UK: London, UK.)
- Symonds, M. R. E., Gemmill, N. J., Braisher, T. L., Gorringer, K. L., and Elgar, M. A. (2006). Gender differences in publication output: towards an unbiased metric of research performance. *PLoS One* **1**, e127. doi:10.1371/JOURNAL.PONE.0000127
- Tabachnick, B. G., and Fidell, L. S. (2014). 'Using Multivariate Statistics.' (Pearson Education: Harlow, UK.)
- Toulmin, S., Rieke, R., and Janik, A. (1984). 'An Introduction to Reasoning.' (Macmillan: New York, NY, USA.)
- Underwood, A. J. (1997). 'Experiments in Ecology: their Logical Design and Interpretation Using Analysis of Variance.' (Cambridge University Press: New York, NY, USA.)
- Universities Australia (2017). '2016 Selected Inter-institutional Gender Equity Statistics.' (Universities Australia: Canberra, ACT, Australia.)
- van den Besselaar, P., and Sandström, U. (2016). Gender differences in research performance and its impact on careers: a longitudinal case study. *Scientometrics* **106**, 143–162. doi:10.1007/S11192-015-1775-3
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics* **10**, 365–391. doi:10.1016/J.JOI.2016.02.007
- Ward, H. B., and Whipple, G. C. (1918). 'Fresh-water Biology.' (Wiley: New York, NY, USA.)
- Warner, P. C., and Ewing, M. S. (2002). Wading in the water: women aquatic biologists coping with clothing, 1877–1945. *Bioscience* **52**, 97–104. doi:10.1641/0006-3568(2002)052[0097:WITWWA]2.0.CO;2
- Wendl, M. C. (2007). H-index: however ranked, citations need context. *Nature* **449**, 403. doi:10.1038/449403B
- Wilhite, A. W., and Fong, E. A. (2012). Coercive citation in academic publishing. *Science* **335**, 542–543. doi:10.1126/SCIENCE.1212540

Handling Editor: Rebecca Lester