Update
TRENDS in Ecology and Evolution Vol.21 No.4 April 2006
Full text provided by www.sciencedirect.com

Research Focus

The *h index* and career assessment by numbers

Clint D. Kelly and Michael D. Jennions

School of Botany & Zoology, The Australian National University, Canberra, ACT 0200, Australia

Growing demand to quantify the research output from public funding has tempted funding agencies, promotion committees and employers to treat numerical indices of research output more seriously. So many assessment exercises are now conducted worldwide that traditional peer assessment is threatened. Here, we describe a new citation-based index (Hirsh's *h index*) and examine several factors that might influence it for ecologists and evolutionary biologists, such as gender, country of residence, subdiscipline and total publication output. We suggest that *h* is not obviously superior to other indices that rely on citations and publication counts to assess research performance.

Introduction

Several different citation-based indices are used to measure research performance (e.g. the number of highly cited papers published, the mean number of citations per paper and the total number of citations). There are valid reservations about using these indices to measure performance because papers are cited for reasons that are unrelated to the quality or utility of a study. For example, researchers cite papers more often by potential reviewers or editors [1] or by colleagues from the same country [2] (reviewed in [1]). Comparing researchers using these indices is problematic because citation patterns vary among scientific disciplines [3]. For biologists, there might even be a taxon effect. For example, herpetologists write conceptually broader introductions than do ornithologists, who focus more narrowly on birds [4]. This should generate an asymmetry in taxon-based citations, resulting in higher citation rates for biologists working on 'popular' organisms. It is therefore important to interpret these indices cautiously and to recognize their limitations.

The latest citation-based index

Recently, *Nature* [5] and *Science* [6] promoted a new measure of research performance developed by Jorge Hirsch called the 'h index' [7], defined as the maximum number of papers h by a scientist where each paper has received h or more citations. It can be calculated using a database such as Thomson Scientific's Web of Science© (http://isiknowledge.com) and sorting publications using the 'times cited' option: scroll down the output until the rank of the paper (in terms of citations) is greater than the number of citations that it has. The preceding rank equals h. For example, if an author's 15th most cited paper has been cited 15 or more times but their 16th most cited has been cited (16 times then their h index is 15. Hirsch notes that top physicists have h values of 60–110.

It should be immediately apparent that h depends on scientific age (i.e. years publishing) because the pool of published papers, and the citations that each receives, increases over time. Hirsch modelled the accumulation of citations and papers to argue that h and age have an almost linear relationship although it might not be so for some individuals. Ultimately, the relationship must become non-linear because the maximum h index is the total number of papers published by a researcher. More generally, Hirsch makes the simplifying assumption that papers accumulate citations at a fixed rate (c). It is well known, however, that most papers enjoy a limited period after publication in which they are cited [8]. To compare individuals of different scientific ages, Hirsch divided h by their scientific age to generate the value m (m can be thought of as the speed with which a researcher's h index increases). Hirsch defined 'scientific age' as the number of years since an author's first publication. For top physicists, m is 1.41–3.89.

The alleged advantages that h and m have over other citation-based indices or counting publications is to favour those authors who produce a series of influential papers rather than those authors who either produce many papers that are soon forgotten or produce a few that are uncharacteristically influential. In addition, some indices can be manipulated by researchers. For example, 'citations per paper' can be increased by publishing fewer papers and by not publishing work that is unlikely to be widely cited. 'Total citations' can be increased by preferentially publishing reviews, which are usually cited more often than are primary data papers. By contrast, Hirsch argues that h is difficult to manipulate; thus, using it for assessment might decrease the likelihood that politically astute, but not necessarily outstanding, scientists gain prestige. He argues that h could more objectively reward scientists with promotions, awards or even funding. If a prerequisite h value is set for, say, appointment to a scientific academy, researchers with high m values will be rewarded because they will reach this value at a younger age. One reason to use h rather than m is that m can fluctuate widely early in a researcher's career.

Measuring *h* for ecologists and evolutionary biologists

The rate at which papers accumulate citations varies across disciplines: cell biology publications accumulate citations more rapidly than do ecological publications [3]. The average number of papers per researcher and

Corresponding author: Kelly, C.D. (Clint.Kelly@anu.edu.au).
Available online 3 February 2006

Box 1. Data set and methods

We quantified the publication output of 187 individual Editorial Board members of seven non-taxon-based journals that cover a range of topics in ecology and evolution (Table I). Using the Web of Science© (http://isiknowledge.com), we searched using the researcher's surname and initials and counted the total number of papers published (total publication output). This database covers scientific articles and does not include other publications such as books and technical reports. We then followed Hirsch's method to calculate the *h* index and *m* (see main text). We defined scientific age as the difference in years between 2005 and the first year in which the researcher published two or more papers. Hirsch used the year in which a researcher's first paper was published. We argue that our approach better represents the beginning of sustained publication. We avoided erroneously including papers by other researchers with the same surname and initials. We noted the current address of the author as their place of residence and divided this into five categories: the UK ($N=19$), non-English-speaking EU countries ($N=29$), Canada ($N=20$), the USA ($N=106$), and 'the rest of the world' ($N=13$).

For linear regressions, we found that log-transformations of scientific age, total publications and the *h index* provided a better fit when calculating regressions. This was due to reducing variance heterogeneity rather than obvious non-linearity in regressions that used untransformed values.

Table I. Data sources

Journal (IF[a])	*N*	Sample used[b]
Am. Nat. (4.48)	33	Every second member of Editorial Board
Behav. Ecol. (2.12)	21	Editor-in-Chief, Editors and Editorial Board
Ecology (4.10)	27	Editorial Board whose term ends in 2006
Evolution (3.72)	31	Associate Editors whose terms run through 2005 and 2006
Mol. Ecol. (4.38)	34	Editorial Board
J. Vert. Paleontol. (1.33)	14	Senior and receiving Editors, Editorial Board 2004–2005
Trends Ecol. Evol. (12.94)	27	Advisory Editorial Board (excludes R. May's work on pure physics)

[a]IF, impact factor of journal (number of times papers published in the journal in 2002 and 2003 were cited in 2004.

[b]Four individuals were excluded because their names are shared by many others (e.g. J. Smith) precluding accurate identification of all their publications.

references per paper also vary across disciplines. How does this effect *h*, which depends on the number of papers published and the citations they have at the time of assessment? Hirsch noted that highly cited biological and biomedical scientists, as ranked by Thomson Scientific, have *h* values of 120–197, much higher than those of physicists. So what are representative *h* values for evolutionary biologists and ecologists? Inspection of a sample of 18 evolutionists and ecologists ranked by Thompson Scientific as 'highly cited' yielded a mean *h* of 45.0 ± 11.4SD and *m* of 1.54 ± 0.42SD. However, rather than focus on the absolute elite, we collected data on Editorial Boards from seven evolution and ecology journals (Box 1).

Of course, Editorial Board members are usually successful researchers so we do not claim their *h* values rigorously represent the average researcher. We used Editorial Boards because journals define fields of research, enabling us to generate objectively a pool of research-active scientists. We selected journals covering broad areas in ecology and evolution, but readers with specialized interests could use the same approach (e.g. bird journals to assess *h* for ornithologists). Also, journals encourage diversity in their Editorial Boards so they probably encompass a wider range of personality types and publications strategies than, say, members of Scientific Academies. By contrast, using membership lists of scientific societies is problematic because many are students or part-time researchers. Similarly, using only tenured academics could be misleading, as some are no longer active researchers.

What factors influence *h*?

Although there was a linear relationship between *h* and scientific age, we used a log-log plot because it provided a better fit than the untransformed values (Box 1) (Figure 1a). Importantly, the fit between the residuals from this regression (=age-corrected *h*) and *m* was high ($R^2=80$–94%) for all journals, except *J. Vert. Paleontol.* ($R^2=32\%$). The assumption of linearity and using *m* to control for scientific age are therefore appropriate. In support of the argument that some fields have lower *m* values than do others owing to differences in citation patterns and/or publication rate, the mean value of *m* varied among journals (Figure 1a). It was lowest for *J. Vert. Paleontol.* ($m=0.67$) and *Ecology* ($m=0.96$). It did not, however, differ significantly among the other five journals ($m=1.14$–1.29); thus, we used only these in our subsequent analyses.

Controlling for scientific age, females had lower *h* values than did males (Figure 1b). Why might this be so? There could be discrimination against citing papers with female authors, or females might produce papers that are less citable, or females might publish fewer papers [9] and *h* and *m* depend, in part, on total publication output. There is no evidence that gender affects citation and publication in ecology [1,10] so the first two explanations seem improbable. However, females did publish fewer papers than did males (79 versus 52 papers after 19 years). After controlling for this, there was no gender difference in *m*. There was also no evidence that female Editorial Board members are less-established scientists (scientific age: 19.0 versus 18.6 years; $t_{144}=0.30$, $P=0.77$). This was because the more papers a researcher published per year the higher his or her *m* value (Figure 1c). This pattern is worthy of closer scrutiny. Do more-productive scientists produce inherently more citable papers? Is there a 'fast-food' effect whereby researchers cite papers or authors that they most often encounter or recall rather than searching for the best reference? Or, could it simply be a lottery effect whereby the likelihood of getting a sufficiently highly cited paper to elevate *h* increases the more you publish?

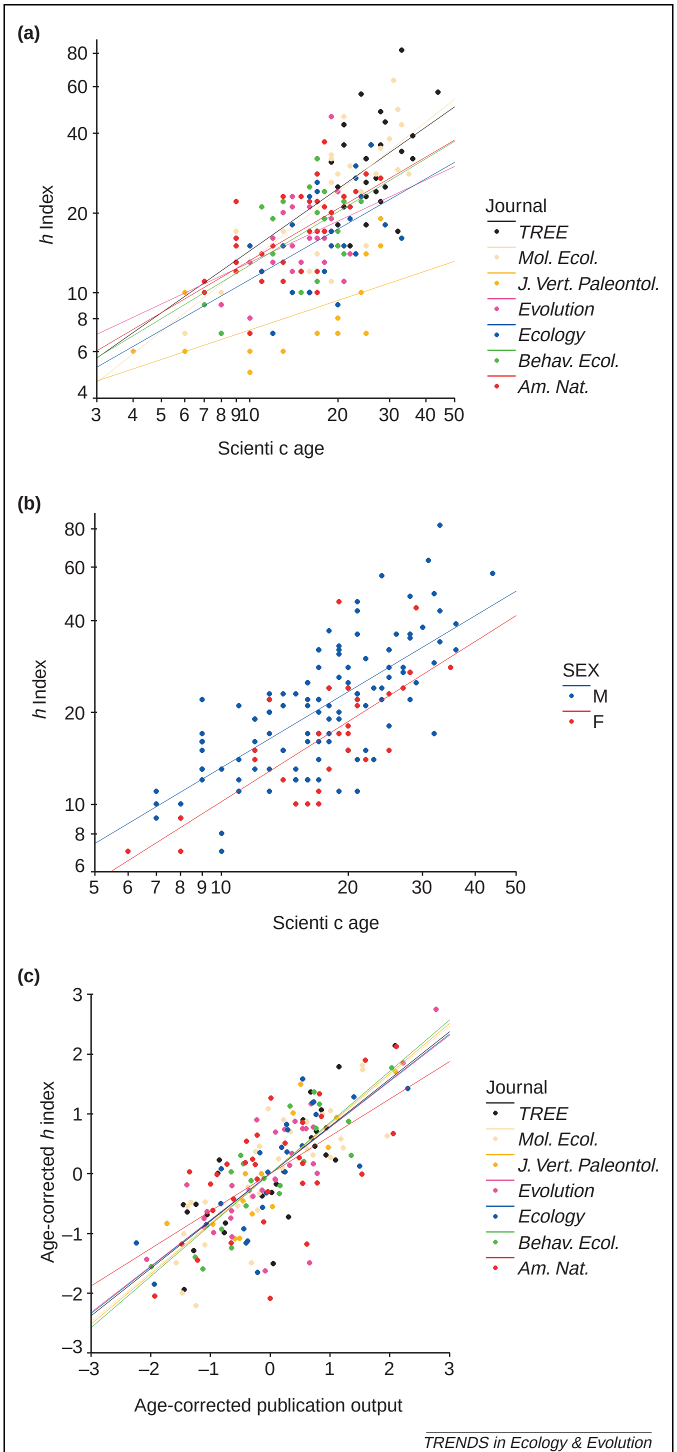

Figure 1. Factors affecting the *h* index. **(a)** There was a positive relationship between log *h* index and log scientific age ($F_{1,179}=87.0$, $P<0.001$) that did not differ among journals. The residuals from this regression (= 'age-corrected *h* index') were closely correlated with *m* (the speed with which a researcher's *h* index increases; $F_{1,185}=434.4$, $P<0.0001$). The mean value of *m* varied significantly among the seven journals ($F_{6,186}=5.2$, $P<0.001$). However, after excluding *J. Vert. Paleontol.* and *Ecology*, whose members had lower *m* values, the mean did not differ among the other five journals ($F_{4,141}=0.65$, $P=0.63$); **(b)** Controlling for age, females had significantly lower *h*-index values than did males ($F_{1,143}=12.2$, $P=0.001$; i.e. $m=1.00\pm0.07$SE, $N=30$ versus 1.27 ± 0.04SE, $N=116$) (excludes *J. Vert. Paleontol.* and *Ecology*); **(c)** Age-corrected publication output (= residuals from the regression of log total publications on log scientific age) and age-corrected *h* index were closely related ($F_{1,144}=205.6$, $P<0.001$). There was no gender difference in *h* index once publication rate was taken into consideration ($F_{1,143}=0.06$, $P=0.81$). The gender effect was attributable to females publishing significantly fewer papers than did males ($F_{1,143}=21.4$, $P<0.001$).

Hirsch [7] argued that self-citation has little effect on *h* values. Is this true? The more papers published the more often someone can self-cite, potentially elevating *h*. Productive scientists often have more collaborators and students with whom they publish, and if these colleagues cite joint publications this could also elevate the focal researcher's *h*. Indirect evidence that self-citation has an effect in ecology is that multi-authored papers are cited more often [1]. Calculating *m* after excluding self-citation is time consuming as each paper citing the focal researcher has to be identified. Inspection of seven randomly selected researchers spanning the range of *h* values suggests that self-citation influences *m*. We recorded a decline from a mean of 1.38 to 1.21 (equivalent to *h* declining from 28 to 24 after 20 years of publishing).

Finally, as place of residence affects citation patterns [1,2], we tested whether it influences *m* and found that it did ($F_{4,139}=3.2$, $P=0.015$). After correcting for gender, the mean *m* for UK residents did not differ from residents of the rest of the EU, but it was greater than that for residents of Canada, the USA or the remaining countries pooled (all pairwise, $P<0.05$). This was not true, however, if Sir Robert May returned to his native Australia ($F_{4,139}=2.1$, $P=0.10$).

Important facts about the *h* index

- The *h index* is closely correlated with total publication output; thus, it will generally result in the same assessment as one based on counting publications. Therefore, if the link between publication rate and *h* is causal, we are unlikely to see researchers shift to producing fewer papers if *h* is widely deployed.
- Female scientists produce fewer papers [9], which affects their *h* index. Assessors should be aware that the *h* index also shows a gender effect.
- Comparison of highly cited scientists suggests *h* and *m* values are lower for evolutionary biologists and ecologists than for biomedical researchers. Administrators and funding bodies should recognize that, as with journal impact factors (IF) [3], the *h* index should not be used to compare the relative importance of researchers in disparate disciplines.
- There is a wide spread in *m* values among Editorial Board members: 90% fell between 0.52 and 1.89 (modal $m=1.07$). A small *m* is therefore poor evidence that a researcher's work is held in low regard by their peers. The journals we used had fairly high IFs (Box 1, Table I) and it needs to be determined whether including specialized journals with lower IFs would reduce *m*. We anecdotally note that *J. Avian Biol.* (IF = 1.66) has a mean *m* of 1.10.

Scientific age is hard to define and $m=h$/age. For younger researchers, a one-year difference in the onset of publishing and a tiny change in *h* can dramatically affect *m* (e.g. 1.60 versus 1.17 for 8/5 versus 7/6).

Finally, what do *h* indices really tell us? Of course, it is distasteful to reduce a lifetime's work to a number. Some scientists make huge contributions through their mentoring and generosity with ideas, skills and time. Without them, academia would collapse: a department solely comprised of relentless publishers would be a joyless place for eager students. Leaving that aside, our impression is that h and m contain small kernels of truth. We think it would be foolhardy to use modest differences in m values to rank individuals, but researchers with high m values, say ≥ 1.5, are those who would, by a conventional peer-review process, also be ranked as highly influential. That said, W.D. Hamilton, E.O. Wilson, R. Trivers, R. Dawkins and S.J. Gould all have $m < 1$, which neatly illustrates the risk of indiscriminate use of the h index.

References

1 Leimu, R. and Koricheva, J. (2005) What determines the citation frequency of ecological papers? *Trends Ecol. Evol.* 20, 28–32
2 Wong, B.B.M. and Kokko, H. (2005) Is science as global as we think? *Trends Ecol. Evol.* 20, 475–476
3 Kokko, H. and Sutherland, W.J. (1999) What do impact factors tell us? *Trends Ecol. Evol.* 14, 382–384
4 Bonnett, X. *et al*. (2002) Taxonomic chauvinism. *Trends Ecol. Evol.* 17, 1–3
5 Ball, P. (2005) Index aims for fair ranking of scientists. *Nature* 436, 900
6 Holden, C. (2005) Random samples: Data point – Impact factor. *Science* 309, 1181
7 Hirsch, J.E. (2005) An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. U. S. A.* 102, 16569–16572
8 Redner, S. (2005) Citation statistics from 110 years of Physical Review. *Phys. Today* 58, 49–54
9 Sax, L.J. *et al*. (2002) Faculty research productivity: exploring the role of gender and family-related factors. *Res. Higher Educ.* 43, 423–446
10 Tregenza, T. (2002) Gender bias in the refereeing process? *Trends Ecol. Evol.* 17, 349–350

doi:10.1016/j.tree.2006.01.005

Female finery is not for males

Natasha R. LeBas

School of Animal Biology, The University of Western Australia, Crawley, WA 6009, Australia

There has recently been an increase in interest in the notion that female ornamentation is selected through male choice, rather than being an artefact of selection on male ornamentation. There are, however, key differences between the sexes in the type of selection pressures that are likely to generate ornamentation and important differences in investment tradeoffs. Here, I discuss that female ornamentation might be selected more often through female competition over resources than through competition over mates, as exemplified in a recent study by Heinsohn and colleagues.

Shifting paradigms

Lower investment in offspring and higher variance in reproductive success usually leads to sexually selected ornamentation in males rather than in females [1]. In accordance, there are numerous studies documenting male competition and female choice of males [1]. By contrast, female ornamentation has been considered an artefact of a genetic correlation with male ornamentation and is rarely investigated [2,3]. This paradigm has recently been questioned [3] and the broader circumstances under which males might be expected to be prudent over mating partners have been recognized [4–7]. In addition, comparative studies suggest that female ornamentation can evolve independently of male ornamentation [3]. Most recent empirical studies of female ornamentation have concentrated on male choice of females and, although there are increasing examples of such choice [3,8–12], they remain rare; in some cases, the benefit to the female from investment in ornamentation appears inadequate to offset the cost of the ornamentation. A recent study by Heinsohn *et al.* [13] highlights a much understudied alternative explanation for female ornamentation, that of female resource competition.

Colourful parrots

Heinsohn and colleagues [13] examined crypsis and intrasexual competition in the highly sexually dichromatic *Eclectus roratus* parrot (Figure 1), which has a polygynandrous mating system (both the male and female have multiple mates) and females reside at their nest tree for ~11 months each year. Up to five males attend each nest and provide all the food for the female and offspring over the breeding period. Females are a vivid blue and red, whereas males are bright green, yet Heinsohn *et al.* state that there is no evidence that *E. roratus* are sex-role reversed. Instead, observations suggest that intrasexual competition is strong in both sexes. Females compete over and defend rare breeding hollows and have been observed to kill one another in these aggressive encounters. Before their restriction to their nest hollow for breeding, females display high in the canopy. Heinsohn *et al.* used sophisticated colour analysis techniques to show that the red and blue coloration of the females gave them the greatest contrast against this leafy background. Given that females nest in hollows, the authors propose that the evolution of their bright coloration has been unconstrained by selection for crypsis during egg incubation.

Corresponding author: LeBas, N.R. (nlebas@cyllene.uwa.edu.au).
Available online 7 February 2006