**References**

1 Restif, O. and Koella, J.C. (2004) Concurrent evolution of resistance and tolerance to pathogens. *Am. Nat.* 164, E90–E102
2 Miller, M.R. *et al.* (2005) The evolution of host resistance: tolerance and control as distinct strategies. *J. Theor. Biol.* 236, 198–207
3 Fineblum, W.L. and Rausher, M.D. (1995) Tradeoff between resistance and tolerance to herbivore damage in a morning glory. *Nature* 377, 517–520
4 Tiffin, P. (2000) Mechanisms of tolerance to herbivore damage: what do we know? *Evol. Ecol.* 14, 523–536
5 Kover, P.X. and Schaal, B.A. (2002) Genetic variation for disease resistance and tolerance among *Arabidopsis thaliana* accessions. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11270–11274
6 Råberg, L. *et al.* (2007) Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. *Science* 318, 812–814
7 Roy, B.A. and Kirchner, J.W. (2000) Evolutionary dynamics of pathogen resistance and tolerance. *Evolution Int. J. Org. Evolution* 54, 51–63
8 Miller, M.R. *et al.* (2006) The evolution of parasites in response to tolerance in their hosts: the good, the bad, and apparent commensalism. *Evolution Int. J. Org. Evolution* 60, 945–956
9 Simms, E.L. (2000) Defining tolerance as a norm of reaction. *Evol. Ecol.* 14, 563–570
10 Mackinnon, M.J. and Read, A.F. (2004) Virulence in malaria: an evolutionary viewpoint. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 359, 965–986
11 Grech, K. *et al.* (2006) Host-parasite interactions for virulence and resistance in a malaria model system. *J. Evol. Biol.* 19, 1620–1630
12 Mauricio, R. *et al.* (1997) Variation in the defense strategies of plants: are resistance and tolerance mutually exclusive? *Ecology* 78, 1301–1311
13 Boots, M. and Haraguchi, Y. (1999) The evolution of costly resistance in host-parasite systems. *Am. Nat.* 153, 359–370
14 Lloyd-Smith, J.O. *et al.* (2005) Superspreading and the effect of individual variation on disease emergence. *Nature* 438, 355–359

0169-5347/$ – see front matter 
doi:10.1016/j.tree.2008.01.006

Letters

# The *h*-index and self-citations

**Leif Engqvist[3] and Joachim G. Frommen**

Department of Evolutionary Biology and Ecology, University of Bonn, An der Immenburg 1, 53121 Bonn, Germany

The *h*-index [1] has been claimed to provide a simple way to compare objectively the scientific achievement of researchers and has rapidly become one of the most favoured measures of scientific output [2]. The *h*-index is an author's number of articles ($h$) that have received at least $h$ citations [1], and thus depends on the number of a researcher's publications and their impact. Some recent articles have called for cautious use of the *h*-index [3–7]. In particular, its robustness against self-citations has been disputed [3–5]. As the enhancement of the *h*-index will often be impeded by the lack of a few citations only, it has been argued that the *h*-index might be susceptible to manipulation by self-citation of such articles [1,4,5]. Here we use simple arguments and quantitative analysis to show that the alleged sensitivity of the *h*-index to self-citations is overestimated, and manipulations thus difficult.

First, it has been claimed that the *h*-index could be inflated through selective citations in the following manner. Typically, a scientist has at least one paper with citations just below the value of the *h*-index. The scientist could now intend to increase the *h*-index by citing this paper [1,4,5]. The impact of self-citation then depends on how often this paper is cited by others. For a well-cited paper, there will only be a temporary effect, because the *h*-index would have increased anyway. If the paper is poorly cited, there will also only be a temporary effect on the *h*-index: because of the skewed distribution of citations [1], there will be other papers with one or two citations below the *h*-value. One of these is likely to be more frequently cited than the poorly cited one, and will therefore soon enough outdistance the self-cited paper. Thus, the only way for the scientist to induce a lasting increase in the *h*-index by selective self-citation is to repeatedly cite a paper at the *h*-index borderline, which would cause an *h*-index that is self-inflated by one. Additionally, the focal paper must be cited continuously at the same rate as the increase rate of the *h*-index. A slower rate would cause the paper to drop below the critical citation rate, and any self-citations would be irrelevant for the *h*-index. Selective self-citations are thus unlikely to have more than a negligible effect on an author's *h*-index. Moreover, it will take considerable time between citing a paper in a newly submitted manuscript and its appearance on the Web of Science. It will be hard to pinpoint the crucial paper for the enhancement of the *h*-index several months in advance.

To examine whether these arguments hold true for a critical assessment, we performed a literature study, selecting 40 authors from the fields of evolutionary biology and ecology (Figure 1) and identified the citation causing their most recent increases in $h$. Next, we distinguished the first citation appearing thereafter, which would have caused the same increase in the author's $h$. The difference between the publication dates of these two citations gives the time that the *h*-index is dependent on one single citation. This time measure thus gives an estimate of how long selective self-citation of target papers will be effective.

More than 20% of all *h*-increasing citations were redundant within a month, 50% within 2 months, and no more than ~10% of all *h*-enhancing citations were crucial for an author's *h*-index after more than 9 months (Figure 1). Nine months is rather an underestimate of the time between the submission of a paper and its registration into the Web of Science. Nevertheless, after that time, there is only a 10% chance that the target citation will actually increase the author's *h*-index. Furthermore, even when the selective self-citation successfully increases the *h*-index, it will rarely be important for more than a few months.

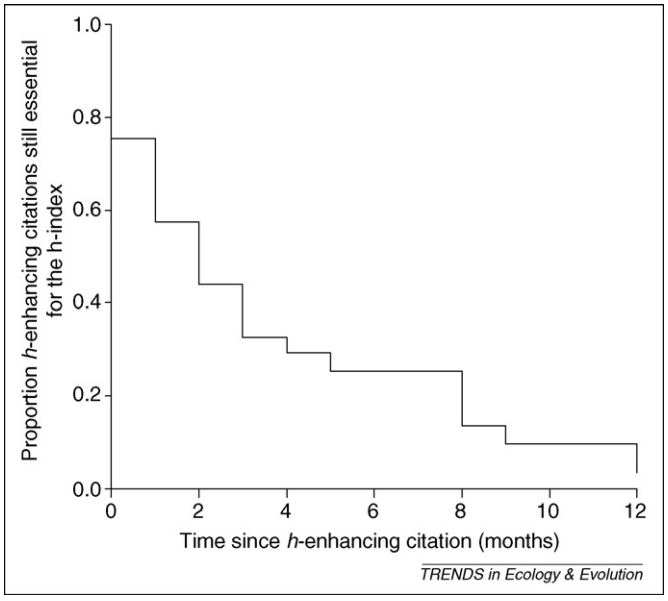



**Figure 1**. A Kaplan-Meier survival plot illustrating the short life expectancy of *h*-enhancing citations for 40 randomly chosen scientists. These were chosen by taking the senior authors of the first ten articles in each of the October 2006 issues of *Trends in Ecology and Evolution*, *Evolution*, *Ecology*, and *Behavioral Ecology*. From the publication record of these scientists registered in the Science Citation Index Expanded database from Web of Science (© Thomson Scientific), we took their respective *h*-indices. We did not consider authors whose *h*-indices could not be precisely determined because they share a common name (e.g. Jane Smith) with other scientists in a similar research field. Furthermore, authors with five or fewer papers were excluded, because here citation rate is often too low to be of any significance. It can be seen that there is usually only a short period where an author's *h*-index depends on a single citation only. Thus, attempts to shrewdly increase the *h*-index by selective self-citations have only a low potential to succeed or have a lasting effect.

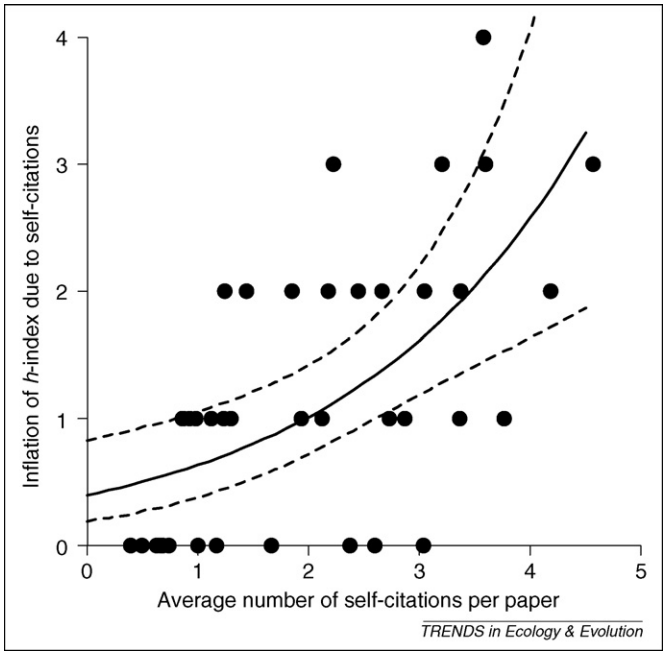



**Figure 2**. The self-citation-induced increase in an author's *h*-index plotted against the author's self-citation frequency (number of self-citations per published paper). The lines give the predicted value ± confidence interval (dashed lines) from a generalised linear model with Poisson errors linked to a log function. This data set was generated in a similar fashion as the previous, with the exception that the October 2007 issues of the respective journals were used. An author's self-citation rate does influence the *h*-index, yet the effect is small. Even excessive self-citation has low potential to increase the *h*-index by more than one unit compared to more restrictive self-citation.

Second, it has been suggested that even random self-citation can inflate the *h*-index. In support of this argument, excluding the self-citations of an author has been shown to result in a decline in *h* [3,5]. Therefore, it has been recommended that the *h*-index should be corrected for self-citations [3,5]. However, as the *h*-index is a citation-based measurement, it is quite obvious that excluding a subset of an author's citations will automatically reduce it. Yet, if all scientists' *h*-indices are inflated by a similar value, self-citations would still not be much of a problem. The question then becomes whether there is systematic variation in the *h*-index that is linked to the author's self-citation rate. We performed a second literature survey to estimate the effect of the total self-citation rate on the *h*-index. For 40 authors, we calculated the decrease in *h* after excluding all self-citations (Figure 2). This decrease was rather modest for all authors (median: one; first quartile: zero; third quartile: two). Only five *h*-indices decreased by more than two.

Counting all self-citations revealed that, on average, each author attained $2.02 \pm 1.13$ self-citations per paper. There was a significant correlation between the frequency of self-citations and the elevation in the *h*-index ($z = 0.47 \pm 0.12$, $\chi^2 = 14.1$, $df = 1$, $p < 0.001$; Figure 2). Thus, by excessive and consistent self-citation, it is possible to increase the *h*-index. However, within the occurring range of self-citation frequency, the effect was reassuringly small. For instance, tripling the number of self-citations from one to three per paper, which would be among the highest rate of self-citation found in this study, would on average elevate the *h*-index by one (Figure 2).

The decrease in the *h*-index when self-citations are removed might to a large extent be caused by the removal of citations *per* se. To demonstrate this, we used the same database as above and excluded the mean self-citation rate (14.4%) for all authors, irrespective of their actual number of self-citations. The decrease in *h* followed a similar distribution as the decrease after excluding true self-citations (Kolmogorov-Smirnov test: $p = 0.98$). The *h*-index reduction when self-citations were excluded correlated significantly with the arbitrary *h*-index reduction (Spearman's rank correlation: $r_s = 0.68$, $n = 40$, $p < 0.001$). Moreover, there was no significant difference in the reduction of an author's *h*-index when random instead of true self-citations were removed (Wilcoxon signed-rank test: $W = 176.5$, $n = 40$, $p = 0.2$). Thus, when self-citations are excluded, some authors' *h*-indices decrease slightly more than others, to some extent because they have a higher frequency of self-citations, but also because they have more papers with citations closer to the *h*-borderline.

In conclusion, even excessive self-citation will cause the *h*-indices of two, in other respects, equivalent scientists to differ only slightly, seldom by more than one unit. Thus, the *h*-index seems robust enough against self-citations, as it would in any case be foolhardy to base grant and personnel decisions on only small deviances in the *h*-index.

**Acknowledgements**
We are most grateful to Theo Bakker, Julia Schwarzer, Katja Bargum and two anonymous referees, who provided useful comments on an earlier draft of this manuscript.

**References**

1 Hirsch, J.E. (2005) An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. U. S. A.* 102, 16569–16572

2 Ball, P. (2007) Achievement index climbs the ranks. *Nature* 448, 737

3 Kelly, C.D. and Jennions, M.D. (2006) The *h* index and career assessment by numbers. *Trends Ecol. Evol.* 21, 167–170

4 Purvis, A. (2006) The *h* index: playing the numbers game. *Trends Ecol. Evol.* 21, 422

5 Schreiber, M. (2007) Self-citation corrections for the Hirsch index. *EPL* 78, 30002

6 Bornmann, L. and Daniel, H.D. (2007) What do we know about the *h* index? *J. Am. Soc. Inf. Sci. Tec.* 58, 1381–1385

7 Lehmann, S. *et al.* (2006) Measures for measures. *Nature* 444, 1003–1004

0169-5347/$ – see front matter 
doi:10.1016/j.tree.2008.01.009 Available online 25 March 2008

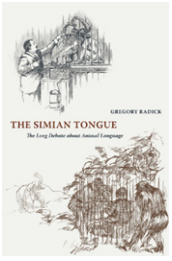
Book Review

# Messages from the Other World

**The Simian Tongue: The Long Debate about Animal Language** by Gregory Radick, University of Chicago Press, 2007. US $45.00 hardcover (544 pages) ISBN 978 0 226 70224 7

## Julia Fischer

Research Group Cognitive Ethology, German Primate Center, 37077 Göttingen, Germany




In mid-September 1892, the self-made scientist Richard Garner set sail for West Africa to lock himself up in a cage in the middle of the Gabon jungle. His quest was to record the utterances of wild chimpanzees and gorillas, and to tear down the language barrier – the assertion that only man had language, while all other creatures lacked it. Garner believed that language had no origins, but existed to a lesser or greater degree in all forms of life. To show just how elaborate the apes' communication was, Garner had even intended to bring one of the latest technological developments into the jungle, an exemplar of Thomas Alva Edison's phonograph. This prompted the comment in the *New York Times* "that only spiritualist circles. . .would receive the phonographic reproductions of the chatter of apes with the same reverent belief that they accord to messages from the other world."

In *The Simian Tongue*, Gregory Radick, a historian of science based at Leeds University, delves deeply into the history of the language origin debate. The book takes its title from one of the early writings of Richard Garner, published in 1891 in the *New Review*. Garner was a celebrity during his time, who combined inquisitiveness, fondness for technical gadgets and a certain sense of adventure, just as Robert Seyfarth and Dorothy Cheney who would many decades later hide reel-to-reel tape recorders in the shade of acacia trees and play back vervet monkey alarm calls to their subjects. Their paper on the semantics of vervet alarm calls published in *Science* [1] was an instant classic. The rise, fall and resurrection of the playback experiment provide the framework for Radick's examination of disputes over the language barrier and the meaning of animal vocalizations.

Radick pays tribute to a large cast of characters involved in this debate. Particularly impressive is the feud between Friedrich Max Müller, a scholar of Sanskrit who had studied in Leipzig, and William Dwight Whitney, based at Yale College. Although Müller maintained that "there is no thought without word, as little as there are words without thought," for Whitney, words were simply signs fixed by convention. Neither tried to be polite to the other. Whitney, for instance, once wrote that Müller's theory of the origin of language "may be summarily dismissed, as wholly unfounded and worthless."

At the beginning of the 20th century, the assertion of the language barrier was fortified by Franz Boas and his fellow cultural anthropologists, who altogether rejected evolutionary accounts of the origin of language. At the same time, the experimental psychologists dominating American campuses showed little interest in the matter. In the third part of the book, Radick thus turns to Europe and the fledgling scientific branch called ethology. He does an impressive job tying in postwar technological developments such as the invention of the sound spectrograph, the rise of information theory and the budding ethological research program. This part greatly profits from extensive interviews with Seyfarth and Cheney as well as the eminent British ethologist Peter Marler, who had directed the Seyfarth research interest toward the question of the meaning of animal vocalizations. Marler also made notes and letters available that give a vivid flavour to the worries and exhilarations of field research. From Marler, for instance, we learn that "now that the *Science* paper is out we will no doubt get a bit of publicity" – a testament to the fact that the importance of media attention is not just a sign of our times.

In *The Simian Tongue*, Radick provides a thorough account of the comparative study of language origins. What is curiously missing from this book, however, is a link to the current debate and today's issues in this field. The book ends with the publication of the vervet monkey alarm call paper, although this publication marks the beginning rather than the end of a highly productive research program. Radick has a love for detail that sometimes left me wondering who would want to know all this stuff (I did). Those with little time on their hands might hence prefer to turn to Radick's earlier, more concise essays [2,3]. *The Simian Tongue* is packed with information: the 400 or so

*Corresponding author:* Fischer, J. (fischer@cog-ethol.de).