EDITORIAL

Impact factor statistics and publication practice: What can we learn?

Michael Taborsky

University of Bern, Switzerland

Abstract

Peer review procedures and citation statistics are important yet often neglected components of the scientific publication process. Here I discuss fundamental consequences of such quality measures for the scientific community and propose three remedial actions: (1) use of a ''Combined Impact Estimate'' as a measure of citation statistics, (2) adoption of an open reviewing policy and (3) acceleration of the publication process in order to raise the reputation of the entire discipline (in our case: behavioural science). Authors, reviewers and editors are invited to contribute to the improvement of publication practice.

In scientific publishing, one measure has developed into a universal remedy. The Journal Citation Reports (JCR) compiled from scientific literature in the ISI database by Thomson Scientific provide a shorthand measure of scientific success. The logic behind this is compelling. Science proceeds through communication. If scientific insight is not noticed, the effort was wasted. If it *is* noticed, it will be cited, or at least that's the idea. The ISI database compiles information about how often a publication was cited, which is a proxy for how often it was read and deemed useful. Indeed, we have an objective measure to separate the wheat from the chaff!

The attentiveness of the scientific audience treats journals differently. Those featuring articles that are often cited score better than others. The measure that has come to be used by the scientific community to judge a journal's quality is the impact factor. It simply divides the number of citations in year 3 to those papers published in the respective journal in the preceding years 1 and 2 by the total number of papers published by the journal in years 1 and 2. This calculation allows comparison of different journals in terms of their conspicuousness: a higher impact factor means that manuscripts published therein receive *on average* more attention.

This is fine and in fact common knowledge, but what about the consequences? They are manifold and important details become obscured. The first consequence is that each journal acquires a tag of quality. The higher the impact factor, the better the journal quality. A direct consequence of this is that ''high impact factor journals'' attract most authors. Who wouldn't want to publish in the best possible organ? Therefore, ''high impact journals'' receive many more submissions than they can print. Selection needs to take effect which ultimately raises the quality of the competing products. 150 years after Darwin this will hardly be dissented. However this selection has two important ramifications: for the publication process, and for those who long for a quick and easy measure of scientific quality.

Let us consider the second ramification first. Most of our research potential relies on taxpayers' money, a precious and limited resource. We, as researchers, need to provide good reasons why the funding we receive is well-earned. If the scientific community regards our output as important, this is a good sign; after all, this community consists of fierce competitors struggling for the few raisins in a tough dough. Therefore, the above-mentioned impact factor statistics seems a godsend as it reflects the importance of a scientific workpiece in the eyes of one's competing peers.

Naturally, there is a caveat. It may take a while until important findings are noticed. So, how should we separate a paper's insignificance from neglect of

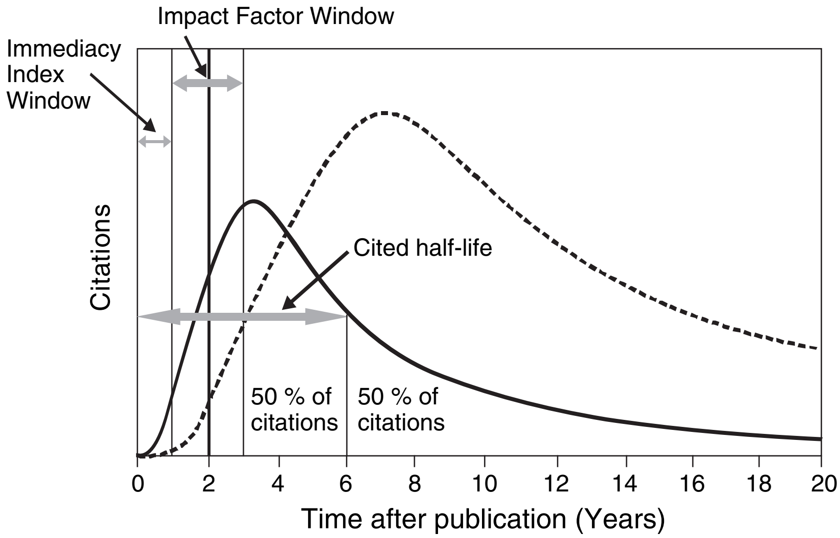

Fig. 1: Illustration of three citation statistics provided by the Journal Citation Reports of the ISI database by Thomson Scientific: impact factor, cited half-life and immediacy index. The solid and dashed lines illustrate generalized citation curves of ''faster'' (but less often cited) and ''slower'' (but more often cited) papers, respectively. See text for further details. Graph drawn by Barbara Taborsky after Amin & Mabe 2003

advertence? Bill Hamilton's influential paper on kin selection published in 1964 was cited four times in the first two years after publication. If one had used this citation figure by the end of 1966 to estimate the importance of his paper, the judgement would have been rather scathing; and very unjust for that matter. By December 2006, this paper was cited 3863 times. Yet obviously we cannot afford to wait fourty something years before making a judgement. So what measure should we ideally use?

Before answering this question, we should take a brief look at the JCR compilation logic. The number of annual citations of a paper is a function changing with time (Fig. 1). It increases until reaching a maximum after some years, before its steady decline. The impact factor of a journal considers the average citation frequency of its publications only in the first two years after appearance. Thereby it introduces an important weighting: Only those papers recognized very quickly will come off well. Those with a more slowly increasing impact will lose out, even if their overall importance is much greater and long-standing (see the dashed line in Fig. 1). A measure for the latter quality is the cited half-life which divides the average annual citation statistic of a journal in half (see Fig. 1). This by itself is again not so useful as a quality assessment because it does not provide a clue as to the absolute height of the citation curve (i.e. the total number of citations received). What we really want to know is the area below the citation curve, i.e. its integral. Unfortunately, this is not provided. A third measure given in the JCR is the ''immediacy index'' which refers to the number of citations to an article in the very year of publication. This is closely linked to the impact factor but even more restricted in time.

A rough approximation to what we really want to know is a combination of two of the measures provided, impact factor (for the height of the curve) and cited half-life (for the skew). By multiplying these two measures we arrive at an estimate closer to the true importance of the product than with both measures taken alone. There would be much better measures than that (e.g. the integral of the citation function), but from what is compiled and publicly available this ''Combined Impact Estimate'' (CIE) is probably the best compromise. A constraint is that the cited half-life statistics are cut off at 10 years in the JCR compilations, so publications with much longer standing are at a disadvantage. Nevertheless, for the sake of pragmatism, we should take what we can get.

What is the effect of such combined measure in comparison to the pure impact factor rating? This can be illustrated with a few examples drawn from the ISI database of 2005. Among ecological journals, Molecular Ecology has an impact factor (IF) of 4.301, thereby ranking among the top ten of the field. The Journal of Animal Ecology (IF 3.399) and Oecologia (IF 3.032) are not quite as highly rated, but their cited half life is more than twice as long as that of Molecular Ecology. Therefore, the combined measure CIE would rank the Journal of Animal Ecology first among these three (CIE 33.99), followed by Oecologia (CIE 29.11) and Molecular Ecology (CIE 20.64). In the field of evolutionary biology, Molecular Phylogenetics and Evolution ranks among the top ten (IF 3.431), while the Biological Journal of the Linnean Society does not reach so high (IF 2.261). Due to its much longer cited half life, it overtakes the molecular journal, however (CIE 17.64 vs. 15.44). In general, the journal rankings are not affected dramatically by this different estimate, but the combined measure adds balance to the comparison between fast moving and long-standing research fields, sub-disciplines and journals.

Editorial procedures are the other important issue affected by the need to strongly select among papers for publication. Usually this involves reviewing by peers and an editorial decision to accept or reject a manuscript based on their comments and recommendations. There is a large body of literature on the pros and cons of peer reviewing – each year about 200 papers are listed within the realm of life sciences alone on this issue by the literature database Medline (Rennie et al. 2003, Berger 2006). Despite enormous effort, scientific scrutiny has not been able to prove consistently that peer reviewing improves the quality of publications in any significant way (Wager & Jefferson 2001, Jefferson et al. 2002, Jefferson et al. 2003, Young 2003, Berger 2006). It has been stressed repeatedly that this process is very slow, expensive, unreliable, open to abuse and frequently unhelpful (e.g. Van Rooyen 1998). It is poor at detecting gross defects and almost useless for detecting fraud, profligate of academic time, while being highly subjective and indeed something of a lottery and very prone to bias (Smith 2006a; see also Nylenna et al. 1994, Godlee et al. 1998). The manuscript quality assessments by different referees hardly coincide (Blank 1991, Howard & Wilkinson 1998). The enormous waste of scientists' time and the absolute, ineluctable bias against innovation have been identified as its worst offences (Roy & Ashburn 2001).

Can we do without peer review? Apparently we could: landmark papers responsible for major advances in science were published without it (see Berger 2006). Still, to question the use of peer review in scientific publishing is regarded by many as outright heresy. With no obvious and promising alternative in sight, presently our aim should be to improve the process rather than to abandon it completely.

The standard practise is that reviewers are anonymous to the authors and that the actor taking decisions about acceptance or rejection, i.e. the editor, is public. However there are important variations on the theme, ranging from complete anonymity among all parties involved to absolutely open procedures and revealed identities. Some journals use double-blind reviewing, where the authors are unknown to the reviewers and vice versa. Advocates of this policy argue that it prevents biased judgement by omitting potential influence of the identity, gender or affiliations of authors. The success of this strategy is limited, however, because only in 53–79% of the cases can author identity be masked to reviewers effectively according to a number of studies across a wide range of disciplines (McNutt et al. 1990, Blank 1991, Cho et al. 1998, Godlee et al. 1998, Justice et al. 1998, van Rooyen et al. 1998, Katz et al. 2002, Snodgrass 2006). Even worse, in cases where the reviewers cannot identify the authors correctly, they might attribute the respective manuscript to ''wrong'' authors, potentially subjecting their response to an even more unjustified bias. Another problem is that the most important players in the game, the editors taking the decision to accept or reject a manuscript, are not blind with regard to the authors which has significant effects (Blank 1991). Moreover, the double-blind reviewing is difficult to administer and does not raise the quality of reviews (Godlee et al. 1998, Justice et al. 1998, van Rooyen et al. 1998), nor does it change the degree to which a review influences editorial decisions (Justice et al. 1998). Since manuscripts of well-known authors are more difficult to mask, and those manuscripts are much more likely to be affected from masking (and hence there the masking would be most important), the inability to mask the identity of well-known authors to reviewers introduces an unfortunate bias and renders this practice dubious (Blank 1991, Justice et al. 1998, Rennie 1998, Godlee 2002; see also Fisher et al. 1994, Young 2003).

Some journals have extended anonymity even to the editorial level. Either the associate editor handling the manuscript or the editor responsible for the ultimate publication decisions are unknown to the authors, even though the authors are of course known to the editors. In such cases, the authors are confined to communicate with an editorial assistant who is not responsible for any decisions on the manuscript. Needless to say, this strategy is not very popular with authors (cf. van Rooyen et al. 1999), partly because scientific argument between author and editor is significantly impeded by this policy and the nasty impression is conveyed to the scientific community that the editor is not prepared to take responsibility for his or her decision. This is of particular concern when assuming that the quality of a journal depends first on good editors (Smith 2006b).

On the bright side, there is an increasing number of journals using open peer review procedures. The aim of the editors of these journals is exactly opposite to dubious collusiveness: complete transparency is advocated for reviewers and authors alike. This comes not without costs, but it has established strengths. Above everything else, this approach has indisputable ethical power. Masked reviewing is a precarious example of privilege and power (that of the reviewer over the fate of the author's manuscript) being dislocated from accountability. In contrast, openness strengthens the link between power and accountability because when reviewers know their names will appear at the end of their reviews, one may be sure that they will be constructive and will attempt to back up their statements (Rennie 1998). Accountability in the reviewing process is essential because it is so important to publish in ''good'' (i.e. high impact) journals, both for the careers of individuals and their research funding (Walsh et al. 2000).

Sceptics of open review put forward that judgement may be less critical, especially if junior scientists are to review work of senior colleagues, for strategic reasons; or that ''old boys'' networks would more easily develop, while on the other side of the coin resentment and animosity between scientists might spread (Fabiato 1994). However, several studies have proven these suspicions ill-founded (see Godlee 2002 for review). Two of four studies testing open peer review against masked forms of reviewing found no difference in review quality, while two found a significant improvement when reviews were signed (McNutt et al. 1990, Godlee et al. 1998, van Royen et al. 1999, Walsh et al. 2000; cf. Smith 2006a). These two studies also found a significant increase in the tendency to recommend manuscript acceptance. In one study the reviewers found that signed reviewing took longer (by 24 minutes) than unsigned reviewing, and in one additional study the procedure took longer (by 25 minutes) when the review was signed and posted on the internet than when it was only signed for the authors (van Royen et al. unpubl., cited in Godlee 2002). One study showed that more referees declined to review when asked to sign their reviews.

So what are the potential costs of open reviewing? (1) Referees may tend to recommend acceptance of a greater proportion of manuscripts than when remaining anonymous, but this appears to hold also when reviewers are blind to authors' identities (Godlee et al. 1998). In any case, this can be easily accounted for by adjusting editorial thresholds. (2) Reviews may take slightly longer, which would translate in more constructive and higher quality reviews, a benefit probably outweighing the costs. (3) Editors may lose a certain proportion of reviewers who are only willing to act in anonymity. Walsh and colleagues (2000) found, however, that the quality of reviews of referees who refused to have their names revealed to authors was significantly lower than the quality of signed reviews (mean = 14%, range = 8.3–20.5%; validated review quality instrument scrutinizing seven items, each scored on a five-point Likert scale; Black et al. 1998). This suggests that the ''loss'' of decliners from an open peer review process may be beneficial to the scientific community in terms of review quality, which should translate into improved publication standards.

A second important benefit of open peer review is that reviewers, who by their altruism spend valuable time free of charge, will receive credit for their substantial effort, even if only from the authors whose work they scrutinized. The flipside of this is the fourth potential cost of open reviewing – and the one probably of most concern to researchers depending on a good publication record: it is the belief that reviewers might elicit resentment and desires of vengeance in colleagues when openly criticizing their manuscript. Bad reputation is the sword of Damocles dangling above the critical (and disclosed) referee. Personally, I think this fear is unfounded. I have signed all reviews I did for dozens of journals in my career, and I frankly admit that more often then not they have been very critical. Still, I do not feel harassed or mistreated by anonymous (!) referees of my own papers, even though a considerable proportion of these referees is most likely the same individuals that were the ''victims'' of my reviewing scrutiny at some stage or another. Signing reviews is no problem as long as they are constructive – and the imperative of constructive criticism in open reviewing is the very merit of this procedure. Reviewers do not make decisions, after all, but give the editor advice, and help improve a manuscript or research effort (cf. Morrison 2006).

When considering this apprehension of retaliation, a risk of masking referees to authors is that they may have a conjecture anyway about the identity of their manuscript's reviewer. In a systematic study this turned out to be true for more than 20% of the reviews involved, and in roughly two thirds of these cases the guess was wrong (Wessely et al. 1996). The latter fact is even more disturbing because it means that an author would more often than not assign a referee's report to a colleague who had nothing to do with it, and the positive or negative feelings involved will implacably hit the wrong target. All this nonsense is avoided by open reviewing.

There are quite a few among us who always sign their reviews, irrespective of a journal's policy, but they are still a minority. The fully unblinded, open review was preferred by only 10% of respondents in a recent survey, mainly on the grounds of transparency and better quality reviews and feedback (Regehr & Bordage 2006; see also Melero & Lopez-Santovena 2001). Other studies found the majority of *authors* to be in favour of open peer review (Van Rooyen et al. 1999). At ETHOLOGY we favour transparency in scientific communication and therefore we do welcome open reviewing by our referees, but no one is obliged to sign their reviews if they do not feel up to it. We prefer such liberal policy presently but may switch to full open reviewing sometime in the future, if this is favoured by a larger proportion of the research community. As Morrison (2006) phrased it, ''a fully open, transparent review process will be the sign of a mature journal reflecting a fair and fully evolved research community.'' Some journals in biology and medicine have adopted this policy (e.g. British Medical Journal (BMJ) and BioMed Central's medical journals (BMC)) and seem to do well with it. Several studies have confirmed that it does work well (e.g. Godlee et al. 1998, van Royen et al. 1998, Walsh et al. 2000). A few journals go even further and publish the reviewers' reports alongside the article on the internet (e.g. Biology Direct or some BioMed Central journals; Wager et al. 2006; see http://www.biomedcentral.com/1472–6904/2/3/prepub for an example) or instigate open reviews *after* publication of the article on the web (e.g. PLoS One). This is a beginning that might ultimately lead to complete transparency in scientific communication. With electronic publishing gradually replacing more traditional forms of scientific dissemination we may face very different publication procedures in the not so far future (cf. Godlee 2002, Smith 2006b).

The new dynamics in scientific publishing should not deceive us about the ever increasing importance of scientific quality assessment by publication statistics. Researchers should be particularly alert about the political and administrative schemes making use of impact factor statistics or similar estimates, even if they are based on a combined and therefore more representative measure such as the CIE outlined above. It is clear to everybody involved in science that any measure based on citation statistics cannot be appropriately compared between research fields. In some fields one publication in three years is a major achievement, while in other fields one paper a month is a must to survive. Still others progress by forms other than peer reviewed papers to disseminate their scientific achievements, such as by books or conference contributions. Nevertheless, when a simple measure is at hand, it will be used to allocate resources, despite better knowledge. Therefore, it does not make much sense to hide behind the commonplace that citation statistics are not suited for interdisciplinary comparisons because they are used all the time, implicitly or explicitly. Research fields with very high impact factors on average benefit from the aura of importance, while other fields progressing at a more steady pace are losing out.

Among scientific disciplines, behaviour is endowed with rather moderate impact factors. What can we do about this? Citation statistics depend on various factors, one of which is the size of a field. The rather negligible crowd of ethologists cannot compete with the immense mass of scientists working in human neurobiology and medicine, for example. When comparing citation statistics of most popular, i.e. most cited, papers in ETHOLOGY with a journal of similar size (i.e. published papers per year), a neurobiological/medical focus, *Neurobiology of Learning and Memory*, this disparity becomes immediately clear. While the six most often cited papers published in 2005 in ETHOLOGY were cited on average 5.17 times by mid December 2006 (Andersson 2005, Heinrich & Bugnyar 2005, Jennings et al. 2005, Johnson 2005, Poisbleau et al. 2005, Setchell and Wickings 2005), the comparable figure for *Neurobiology of Learning and Memory* is 10.83 times (Barros et al. 2005, Canales 2005, Kim & Ragozzino 2005, Kohler & Wehner 2005, Kuhlmann et al. 2005, Lambert et al. 2005), i.e. twice as high. Other sociological factors with crucial influence on impact factor statistics include the number of authors (the greater the mean number of authors per paper, the higher the mean impact factor for a subject area; Amin & Mabe 2003) and the type of articles covered by a journal (e.g. letters, original papers, reviews).

One trait of particular importance for citation statistics in a subject area is, however, the time between submission and publication. This interval determines largely the probability and extent to which a paper can raise the impact factor of all journals in the respective field. For example, if there is a publication delay of two years, a published paper will contain citations that are already three years of age or older. This means that this paper will not contribute anything to the impact factors of the journals in the entire field – its citations simply miss the impact factor window. It is worth stressing here that the speed of publication of any journal in a field feeds back on the entire field, which is particularly important if ''big'' journals are concerned, i.e. those used by a substantial proportion of researchers.

Let me illustrate this with an example (here I shall remain with ETHOLOGY so as not to arouse suspicion of practising unfair competition). The six papers published in 2006 with the shortest interval between first submission and appearance in print (mean: 7.66 months) cited on average 6 publications from 2004 or younger (= median; range: 2–11; Blackledge & Zevenbergen, Blumstein 2006, Ferkau & Fischer 2006, Monclus et al. 2006, Safi et al. 2006, Tan & Tang 2006). In contrast, only 3 publications on average (= median; range: 0–4) were cited from 2004 or younger by the six papers with the longest interval between first submission and appearance in print (mean: 18 months; Amo et al. 2006, Bellemain et al. 2006, Friedl 2006, Kutsukake 2006, Peters & Despland 2006, Roper & Zann 2006). This twofold difference is merely a result of differences in the interval between first submission and publication of these papers, which has significant consequences for the JCR statistics and hence impact factors of the entire field.

What can we do about this? Journal editors can try to speed up publication procedures. My plea to all contenders in our discipline is to work hard towards this end because the important issue regarding the use of impact factor statistics is not the differences between journals *within* the field, as these tend to be relatively small and hence are strongly affected by random variation (impact factor differences of moderately sized journals would need to exceed ca. 25% to reach statistical significance; Amin & Mabe 2003). The importance of speed is that *the entire field* will benefit if the average impact factors rise, despite the flaw in the logic of comparisons between disciplines (e.g. behavioural science against neurobiological or molecular sciences). In our attempt to work towards this end, ETHOLOGY has reduced the time between original submission and decision to an average period of 47 days in 2006 (source: manuscript handling statistics from the electronic editorial office Manuscript Central™). Regarding publication times, through our increase of journal volume by nearly one half (cf. Taborsky 2006) we have by now greatly reduced the backlog of papers that had accumulated because of increasing submission numbers. Therefore, publication times have shortened, and the electronic OnlineEarly publication of complete papers ready to print further helps to provide rapid access to papers submitted to this journal.

Authors can also help their research field by their choice of where to publish. If we accept that research fields are judged by ISI citation statistics – and we can do little about this I'm afraid – authors can base their decision as to where to publish on a journal's average time to publication. Not the impact factor of a journal is primarily important for the effect of its publications on the development of citation statistics in that field, but the average publication delays. Authors choosing journals on the basis of this criterion will not only benefit from the earlier dissemination of their work, they will also help to raise the reputation of the entire research field. Unfortunately, it is sometimes difficult to discover the relevant information, even if it regards the most important measure: the time between first submission and publication of the print version. Some journals do not provide such figures at all, while others give the submission date of the *final* instead of the original version (without making this clear), which is highly misleading and - one is tempted to say - even treacherous. This is probably another unfortunate effect of the rat race for the most competitive statistics.

Last but not least, impact factor statistics are influenced by the accessibility of the articles published in a journal. This works in favour of an open access policy, which is now encouraged by Blackwell Publishing through their Online Open option for authors (see: http://www.blackwellpublishing.com/static/onlineopen.asp). Other measures can help as well: most researchers access journals through (electronic) libraries of the institutions to which they are affiliated. Nowadays most libraries have contracts with publishers that place many if not most or all of their journals at the disposal of their clients. Therefore, the more journals there are in a publisher's portfolio, the more likely a scientist will benefit from this circumstance. In this light, the merger between Blackwell Publishing and John Wiley & Sons Inc. that just made this publisher one of the biggest in science (with approximately 1,350 scholarly peer-reviewed journals) is good news for the usability of its journals, and for the authors and readers of ETHOLOGY alike.

Comments welcome: Your comments on this article are welcome and may be published in the journal as a ''letter to the editor''.

References

Amin, M. & Mabe, M. A. 2003: Impact factors: use and abuse. Medicina-Buenos Aires. **63**, 347—354.

Amo, L., Lopez, P. & Martin, J. 2006: Can wall lizards combine chemical and visual cues to discriminate predatory from non-predatory snakes inside refuges? Ethology. **112,** 478—484.

Andersson, M. 2005: Evolution of classical polyandry: Three steps to female emancipation. Ethology. **111**, 1—23.

Barros, D. M., Ramirez, M. R. & Izquierdo, I. 2005: Modulation of working, short- and long-term memory by nicotinic receptors in the basolateral amygdala in rats. Neurobiology of Learning and Memory. **83**, 113—118.

Bellemain, E., Swenson, J. E. & Taberlet, P. 2006: Mating strategies in relation to sexually selected infanticide in a non-social carnivore: The brown bear. Ethology. **112**, 238—246.

Berger, E. 2006: Peer review: A castle built on sand or the bedrock of scientific publishing? Annals of Emergency Medicine. **47**, 157—159.

Black, N., van Rooyen, S., Godlee, F., Smith, R. & Evans, S. 1998: What makes a good reviewer and a good review for a general medical journal? Jama-Journal of the American Medical Association. **280**, 231—233.

Blackledge, T. A. & Zevenbergen, J. M. 2006: Mesh width influences prey retention in spider orb webs. Ethology. **112**, 1194—1201.

Blank, R. M. 1991: The effects of double-blind versus single-blind reviewing: experimental evidence from the American Economic Review. American Economic Review. **81**, 1041—1067.

Blumstein, D. T. 2006: The multipredator hypothesis and the evolutionary persistence of antipredator behavior. Ethology. **112**, 209—217.

Canales, J. J. 2005: Stimulant-induced adaptations in neostriatal matrix and striosome systems: Transiting from instrumental responding to habitual behavior in drug addiction. Neurobiology of Learning and Memory. **83**, 93—103.

Cho, M. K., Justice, A. C., Winker, M. A., Berlin, J. A., Waeckerle, J. F., Callaham, M. L. & Rennie, D. 1998: Masking author identity in peer review - What factors influence masking success? Jama-Journal of the American Medical Association. **280,** 243—245.

Fabiato, A. 1994: Anonymity of reviewers. Cardiovascular Research. **28**, 1134—1139.

Ferkau, C. & Fischer, K. 2006: Costs of reproduction in male Bicyclus anynana and Pieris napi butterflies: Effects of mating history and food limitation. Ethology. **112**, 1117—1127.

Fisher, M., Friedman, S. B. & Strauss, B. 1994: The effects of blinding on acceptance of research papers by peer-review. Jama-Journal of the American Medical Association. **272**, 143—146.

Friedl, T. W. P. 2006: Individual male calling pattern and male mating success in the European treefrog (Hyla arborea): Is there evidence for directional or stabilizing selection on male calling behaviour? Ethology. **112**, 116—126.

Gherardi, F. & Atema, J. 2005: Memory of social partners in hermit crab dominance. Ethology. **111**, 271—285.

Godlee, F., Gale, C. R. & Martyn, C. N. 1998: Effect on the quality of peer review of blinding reviewers and asking them to sign their reports - A randomized controlled trial. Jama-Journal of the American Medical Association. **280**, 237—240.

Godlee, F. 2002: Making reviewers visible - Openness, accountability, and credit. Jama-Journal of the American Medical Association. **287**, 2762—2765.

Hamilton, W. D. 1964: Genetical evolution of social behaviour I. Journal of Theoretical Biology. **7**, 1.

Heinrich, B. & Bugnyar, T. 2005: Testing problem solving in ravens: String-pulling to reach food. Ethology. **111**, 962—976.

Howard, L. & Wilkinson, G. 1998: Peer review and editorial decision-making. British Journal of Psychiatry. **173**, 110—113.

Jefferson, T., Alderson, P., Wager, E. & Davidoff, F. 2002: Effects of editorial peer review - A systematic review. Jama-Journal of the American Medical Association. **287**, 2784—2786.

Jefferson, T. O., Alderson, P., Davidoff, F. & Wager, E. 2003: Editorial peer review for improving the quality of reports of biomedical study (Cochrane Methodology Review). The Cochrane Library (Issue 1). Update Software, Oxford.

Jennings, D. J., Gammell, M. P., Payne, R. J. H. & Hayden, T. J. 2005: An investigation of assessment games during fallow deer fights. Ethology. **111**, 511—525.

Johnson, J. C. 2005: The role of body size in mating interactions of the sexually cannibalistic fishing spider Dolomedes triton. Ethology. **111**, 51—61.

Justice, A. C., Cho, M. K., Winker, M. A., Berlin, J. A. & Rennie, D. 1998: Does masking author identity improve peer review quality? - A randomized controlled trial. Jama-Journal of the American Medical Association. **280**, 240—242.

Katz, D. S., Proto, A. V. & Olmsted, W. W. 2002: Incidence and nature of unblinding by authors: Our experience at two radiology journals with double-blinded peer review policies. American Journal of Roentgenology. **179**, 1415—1417.

Kim, J. & Ragozzino, M. E. 2005: The involvement of the orbitofrontal cortex in learning under changing task contingencies. Neurobiology of Learning and Memory. **83**, 125—133.

Kohler, M. & Wehner, R. 2005: Idiosyncratic route-based memories in desert ants, Melophorus bagoti: How do they interact with path-integration vectors? Neurobiology of Learning and Memory. **83**, 1—12.

Kuhlmann, S., Kirschbaum, C. & Wolf, O. T. 2005: Effects of oral cortisol treatment in healthy young women on memory retrieval of negative and neutral words. Neurobiology of Learning and Memory. **83**, 158—162.
Kutsukake, N. 2006: The context and quality of social relationships affect vigilance behaviour in wild chimpanzees. Ethology. **112**, 581—591.
Lambert, T. J., Fernandez, S. M. & Frick, K. M. 2005: Different types of environmental enrichment have discrepant effects on spatial memory and synaptophysin levels in female mice. Neurobiology of Learning and Memory. **83**, 206—216.
McNutt, R. A., Evans, A. T., Fletcher, R. H. & Fletcher, S. W. 1990: The effects of blinding on the quality of peer-review: a randomized trial. Jama-Journal of the American Medical Association. **263**, 1371—1376.
Melero, R. & Lopez-Santovena, F. 2001: Referees' attitudes toward open peer review and electronic transmission of papers. Food Science and Technology International. **7**, 521—527.
Monclus, R., Rodel, H. G. & von Holst, D. 2006: Fox odour increases vigilance in european rabbits: A study under semi-natural conditions. Ethology. **112**, 1186—1193.
Morrison, J. 2006: The case for open peer review. Medical Education. **40**, 830—831.
Nylenna, M., Riis, P. & Karlsson, Y. 1994: Multiple blinded reviews of the two manuscripts: Effects of referee characteristics and publication language. Jama-Journal of the American Medical Association. **272**, 149—151.
Peters, M. I. & Despland, E. 2006: Plasticity in forest tent caterpillar collective foraging schedules. Ethology. **112**, 521—528.
Poisbleau, M., Fritz, H., Guillemain, M. & Lacroix, A. 2005: Testosterone and linear social dominance status in captive male dabbling ducks in winter. Ethology. **111**, 493—509.
Regehr, G. & Bordage, G. 2006: To blind or not to blind? What authors and reviewers prefer. Medical Education. **40**, 832—839.
Rennie, D. 1998: Freedom and responsibility in medical publication - Setting the balance right. Jama-Journal of the American Medical Association. **280**, 300—302.
Rennie, D., Flanagin, A., Smith, R. & Smith, J. 2003: Fifth international congress on peer review and biomedical publication - Call for research. British Medical Journal. **326**, 563—564.
Roper, A. & Zann, R. 2006: The onset of song learning and song tutor selection in fledgling zebra finches. Ethology. **112**, 458—470.
Roy, R. & Ashburn, J. R. 2001: The perils of peer review. Nature. **414**, 393—394.
Safi, K., Heinzle, J. & Reinhold, K. 2006: Species recognition influences female mate preferences in the common european grasshopper (Chorthippus biguttulus Linnaeus, 1758). Ethology. **112**, 1225—1230.
Setchell, J. M. & Wickings, E. J. 2005: Dominance, status signals and coloration in male mandrills (Mandrillus sphinx). Ethology. **111**, 25—50.
Smith, R. 1999: Opening up BMJ peer review - A beginning that should lead to complete transparency. British Medical Journal. **318**, 4—5.
Smith, R. 2006: (a) Peer review: a flawed process at the heart of science and journals. Journal of the Royal Society of Medicine. **99**, 178—182.
Smith, R. 2006: (b) The highly profitable but unethical business of publishing medical research. Journal of the Royal Society of Medicine. **99**, 452—456.
Snodgrass, R. 2006: Single- versus double-blind reviewing: An analysis of the literature. Sigmod Record. **35**, 8—21.
Taborsky, M. 2006: Ethology into a new era. Ethology. **112**, 1—6.
Tan, E. J. & Tang, B. L. 2006: Looking for food: Molecular neuroethology of invertebrate feeding behavior. Ethology. **112**, 826—832.
van Rooyen, S. 1998: A critical examination of the peer review process. Learned Publishing. **11**, 185—191.
van Rooyen, S., Godlee, F., Evans, S., Smith, R. & Black, N. 1998: Effect of blinding and unmasking on the quality of peer review - A randomized trial. Jama-Journal of the American Medical Association. **280**, 234—237.
van Rooyen, S., Godlee, F., Evans, S., Black, N. & Smith, R. 1999: Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial. British Medical Journal. **318**, 23—27.
Wager, E. & Jefferson, T. 2001: Shortcomings of peer review in biomedical journals. Learned Publishing. **14**, 257—263.
Wager, E., Parkin, E. C. & Tamber, P. S. 2006: Are reviewers suggested by authors as good as those chosen by editors? Results of a rater-blinded, retrospective study. BMC Medicine. **4**, 1—5.
Walsh, E., Rooney, M., Appleby, L. & Wilkinson, G. 2000: Open peer review: a randomised controlled trial. British Journal of Psychiatry. **176**, 47—51.
Wessely, S., Brugha, T., Cowen, P., Smith, L. & Paykel, S. 1996: Do authors know who refereed their paper? A questionnaire survey. British Medical Journal. **313**, 1185.
Young, S. N. 2003: Peer review of manuscripts: theory and practice. Journal of Psychiatry & Neuroscience. **28**, 327—330.

Ethology **113** (2007) 1–8