

Michael Nentwich

Quality control in academic publishing: challenges in the age of cyberscience

Published online: 23 July 2004
© Springer-Verlag 2004

Abstract This article discusses the future of quality control in an academic publication system that will be largely based on electronic publishing. Information and communication technologies both challenge traditional ways and open remedies for existent problems of present gate-keeping. New forms of ex-ante and of ex-post quality control may partly replace and partly amend peer review, citation indices and quality filters based on the reputation of the publisher. Open peer review, online commenting, rating, access counts and use tracking are evaluated and put in perspective. Refuting the common argument that e-publishing leads to less quality, this paper puts forward scenarios of the future quality control system. Most likely, we shall see mixed systems, combining old and new elements, of different shapes in the different research fields.

Zusammenfassung In diesem Artikel wird die Zukunft der Qualitätskontrolle in einem weitgehend auf Elektronisch Publizieren basierenden Veröffentlichungswesen diskutiert. Informations- und Kommunikationstechnologien stellen eine Herausforderung für die traditionellen Verfahren dar und zugleich eine Chance, Probleme des bestehenden Kontrollsystems zu lösen. Neue Formen einer ex-ante und ex-post Qualitätskontrolle könnten das auf der Reputation des Verlags beruhende System von Peer-Review, Referenzindizes und Qualitätsfilter teils ersetzen, teils verbessern. Offene Begutachtung, Online-Kommentierung, Rating, Zugriffszählung und Nutzungsverfolgung werden analysiert und in Perspektive gesetzt. In Widerspruch zu dem verbreiteten Argument, elektronisches Publizieren führe zu einem Qualitätsverlust, werden in diesem Beitrag Szenarien

M. Nentwich
Institut für Technikfolgen-Abschätzung,
Österreichische Akademie der Wissenschaften,
Strohgasse 45/5, 1030 Vienna, Austria
E-mail: mnent@oeaw.ac.at
Tel.: +43-1-51581-6583
Fax: +43-1-710-9883

für ein zukünftiges Qualitätskontrollsystem vorgestellt. Am wahrscheinlichsten werden sich gemischte Systeme entwickeln, in denen sich alte und neue Elemente verbinden, in je unterschiedlicher Gestalt für verschiedene Forschungsgebiete.

Résumé Le présent article se penche sur l'avenir du contrôle de qualité dans un système d'édition universitaire qui sera largement fondé sur la publication électronique. Les technologies de l'information et de la communication lancent un défi aux méthodes traditionnelles, mais ouvrent aussi la voie à la solution des problèmes existants actuellement dans le domaine des systèmes à accès indirect. De nouvelles formes de contrôle de qualité antérieur et postérieur pourront partiellement remplacer et modifier le système de contrôle par les pairs, des index de citations et des filtres de qualité basé sur la réputation de la maison d'édition. Un contrôle ouvert par les pairs, les commentaires en ligne, les évaluations, le comptage des consultations et la traçabilité des utilisateurs sont analysés et mis en perspective. Contrairement à l'argument courant selon lequel la publication électronique entraîne une baisse de qualité, cet article présente des scénarios pour le futur système de contrôle de la qualité. Il est très probable que se développeront des systèmes mixtes combinant des éléments anciens et nouveaux, sous différentes formes pour les différents domaines de recherche.

1 Introduction

No doubt, the use of information and communication technologies has changed science and research. Academics exchange e-mails, participate in online debates, co-operate from a distance, use remote databases, simulate and model realities on their computers, and teach their students on the Web. One of the most visible impacts of the evolution of "cyberscience" (Nentwich 2003) relates to the scholarly publication system. Academic publishing is no longer the same as it was before the advent of the Internet: In many fields, electronic journals emerged, the publishing houses also offer their paper journals online, huge digital working paper archives give access to the research literature at an early stage, and research libraries slowly turn into "cybraries" providing access to digital repositories of all sorts. Furthermore, we can observe new forms of scholarly publications that would not have been possible in the traditional paper environment, but can only be produced in digital formats. Hypertexts present knowledge differently, multimedia enhances the ways to convey messages to the reader, and communicating research results via databases becomes ever more common. In a nutshell, how researchers communicate formally is fundamentally challenged by the new technological options. One of the most widely debated challenges here is quality.

While the established procedures of quality control in the scholarly publication system are a centrepiece of today's research system, e-publishing is affecting these traditions. On the one hand, new technology gives way to new forms of refereeing (in particular open peer commentary and rating systems), which challenge the present system. On the other hand, self-publishing has led to a massive amount of papers available over the net without much, or even any,

quality control. In this paper, I shall present these developments in the age of cyberscience (2), and assess the future of quality control as a cornerstone of academia (3). Inter alia, we shall analyse the widespread argument that a higher quantity of publications, as enabled through cyberscience, would necessarily lead to lower quality (Sect. 3.1), and then discuss implementation issues (Sect. 3.3).¹

2 Cyberscience developments and opportunities

In addition to the many traditional variants of academic quality control (in particular, peer refereeing), there are a number of new opportunities, which are only possible in the age of cyberscience and which may eventually lead to a new system of quality control.

2.1 New forms of ex-ante quality control

The communicative potential of the Internet, in particular WWW publishing combined with e-mail and Web mailing, enables open peer commentary, either in a pure variant (1), or mixed with traditional review systems (2). Furthermore, it allows for differentiated refereeing standards (3).

(1) Open peer review/commentary While traditionally, the research community at large only finds out about the submission once it is printed, open forms of reviewing in e-journals announce that a paper has been submitted. The submission is uploaded to the journal's server and everyone can have a look at it. In the case of revisions (along the path from the original submission to the final published text) the interim steps are also visible. In addition, the editor does not select referees, but opens a general, open debate about the manuscript. Self-appointed referees (Sumner and Shum 1997) comment on the merits of the paper online. In general, both the author's and the public referees' identities are not disguised. While there are good arguments in favour of anonymity, the primary argument against it is that it seems wrong for somebody to make an important judgement on the work of others in secret (Smith 1999). Although, in principle, the technology allows for anonymity, it is not implemented because it would not suit well the otherwise open character of the procedure.

The philosophy behind open peer review is basically that it is expected that the more people participate in the review process, the better the result. An open, non-anonymous procedure would be more "democratic", less determined by a single editor or a small board. In addition, the author is not only at the receiving end of a decision already taken, but in a position to participate in the review process through interactive means. While reviewing in the traditional model is a solitary activity, the open communicative process may lead to a discourse in which the ideas are refined and shaped.

However, open peer commentary is not without problems. Similar to the ex-post variant discussed in Sect. 2.2, the following four dangers exist:

¹This article is a condensed version of Ch. 6 from Nentwich (2003, pp. 367–397)

- (a) Many articles will not get any comments at all, in particular those that are flawed;
- (b) There might be strong competition to attract any form of positive comment. Both (a) and (b) “might accentuate already current tendencies towards scholarly cliquishness”) (Tomlins 1998, p. 138);
- (c) It may be problematic to assure that the self-appointed referees or commentators are qualified specialists (Harnad 1998a). Will the top specialists engage in this extra work at all? In other words: will it be peers, i.e. experts of at least the same level as the authors, who will comment? Perhaps only those at the periphery, who have no other possibility to be taken notice of, will participate; and
- (d) The general arguments in favour of anonymous refereeing are pertinent here as well (in particular, that the reviewer can actually criticise openly without the danger that the author would know the originator and be resentful).

Given these problems, open peer commentary is rather present in mixed systems as a powerful and important supplement to peer review, but not as a substitute for it.

(2) *Mixed systems* With a view to combining the strengths of both worlds of traditional and open peer review, a number of mixed models are being tried out by editors of innovative journals. One alternative borrows the idea of openness and transparency and applies it to the refereeing process. For instance, referees' comments could be made available for open exploration and interaction (Davenport and Cronin 1990). Referee reports could be revised, published and linked to the published article as commentaries if the referee wishes, alongside with author rebuttals as well as other commentaries (Harnad 1998a).

Another alternative combines anonymous reviewing with open peer review at a later stage. Hence there are different participants in the different stages of the whole process. This is practised, for instance, in the Journal of Interactive Media in Education (JIME). Its sophisticated refereeing system (Sumner and Shum 1997) consists of two reviewing phases, first a closed one in which only appointed referees discuss the merits of the manuscript in a dedicated review Web site, and an open one in which the pre-print is discussed publicly. Only after a couple of weeks of online commentary and debate, the paper is finally released.

Other mixed systems could also include two (or more) review phases, one open, the next restricted, opposite to JIME: In this case, the open review process would precede the appointed referee or the editorial board review phase (Odlitzko 1994). The public comments would inform the referees and the editor. A proposal with even multiple open peer commentary phases was also recorded by an OECD report (1998).

(3) *Differentiated refereeing standards* The digital publishing environment allows for various levels of quality standards. Two proposals have been put forward with a view to having explicitly differentiated refereeing standards. First, in the context of a scenario in which academic communications will be hypertexts, refereeing standards can be defined as a function of the module (Kircz and Roosendaal 1996). While at present, articles are peer-reviewed without any discrimination between the various kinds of information in them, in a “world of

well-defined modules, the refereeing standard for a module Method will be distinctly different from the module Data-acquisition” (Kircz 2001). Different experts may also referee different parts. This fragmentation should, however, have its limit. In order to guarantee, nevertheless, an appropriate integrative appraisal of the whole work, taking due account of the interconnection of all parts, an additional supra-referee would have to be appointed who overlooks and integrates the partial assessments. Second, electronic archives could provide the opportunity to publish scientific content that would not be accepted in a traditional journal, not because it does not meet the required quality standards, but because it does not fit the “line” of a journal or simply because it is too long. For example, it was proposed that BiomedCentral journals should also include “the description of experimental work, from both laboratory and clinical investigators, that lacks definitive or ‘positive’ results” (Varmus et al. 1999). While such material is unlikely to be accepted for publication in most current journals, it would often be useful to others contemplating similar experimental approaches. In both cases, quality would be controlled, but not at equal levels to allow for access to innovative or otherwise useful material.

2.2 New forms of ex-post quality control

In the traditional paper-based academic publishing environment, only ex-ante quality control is feasible. If something turns out to be erroneous or needs qualification, it nevertheless stays as part of the published “body of knowledge”. There is no way to delete ex-post a text printed and distributed around the world, nor would it possible to tag the contribution saying that the content “should be handled with care”. Cyberscience, instead, provides for a variety of ex-post mechanisms that could do exactly this.

As a first alternative, I should mention that already today (and even more so in the future), most researchers access publications not directly through browsing the shelves of a library, but by scanning the results of an online query in a bibliographic database. One simple, but dangerous way of deflecting from false results—no matter how this might have been established—would be to remove the item from the “first layer” of the database. The effect of this would be that the item would no longer be found (Sect. 2.3). Other alternatives possible in the digital environment leave the publication and the bibliographic entry untouched, but let readers (as well as the author) attach comments of various sorts. Subsequent readers may use this added information for their own assessment of the content. In the following, I shall discuss three such novel mechanisms: annotation (1), rating (2), and use tracking (3).

(1) Online (ex-post) commenting/annotation Quality control might be provided after publication by reader comments (or annotations) that the new technology could “attach” to any given article. This may be organised for self-published manuscripts as well as for articles formally published in e-journals (whatever the original quality check). The necessary technology is already developed and various experiments are under way. In the Semantic Web initiative, the WWW Consortium has taken up this idea in its ambitious Annotea project. The idea is that positive comments would probably make an article more significant, while negative comments (or no comments) would probably encourage its marginal-

isation. LaPorte et al. (1995) exemplify this with the plan of a “global health information server” where papers would get a public comment card. “If papers are poor, then the scientific community will most certainly indicate that they are poor; this is the nature of science, and this is the nature of the internet” (for law, see Hibbitts 1996; Tomlins 1998). Each review would rate an article without (as in the current system) forcing its (perhaps premature) suppression due to a negative verdict of the reviewers. In such a system, each reader can have the critical warrant that has been the prerogative of the traditional referee (Dav-enport and Cronin 1990).

The comments are stored together with (or at least directly linked to) the paper which becomes a “living document” (Odlyzko 1994). This could be combined with a system allowing the author to submit, from time to time, a new version of the paper including or taking due account of the comments.

The subversive potential of ex-post commenting challenges how the questions of “what is scholarship and who is a scholar” (Nathenson 2001, p. 3) are decided today. Whether the quality of the collective assessment—which certainly includes evaluations of people with insufficient knowledge—meets the quality of appointed referees is an open question. It reminds us of the debate about democracy and whether, for example, a plebiscite can be a valid alternative to expert decisions. This is not only an ideological question. What we can learn from this parallel discussion regarding the political sphere is, *inter alia*, that minimum thresholds may play a role. Similar to political votes which may not be valid because too few people participated in the voting, one could think of designing ex-post quality control systems that do show comments only as soon as a minimum number of people have registered their comments. This is, however, a very tricky issue and needs in-depth pondering and testing. A related issue to consider is that commenting may lead to what may be called “mainstreaming”. While at the same time enabling dissenting opinions to get published at all, “fishing” for positive comments and ratings may tend to favour less controversial contributions. Again, the effect is possible, but difficult to be evaluated in advance without practical experience.

(2) *Rating/scoring* Commenting is but one form of reacting to published manuscripts. It has two “disadvantages”. First, commentators have to write new text which takes time and needs more in-depth reflection; second, subsequent readers, too, need time for orientation as they have to browse through all comments with a view to making sense of all of them before being in a position to decide whether they should go on reading the main article. Either in combination with commenting or independently, “rating” (scoring) has been proposed to solve these two problems. Rating is giving marks by readers (Alton-Scheidl et al. 1997). The marks are collected and means are computed. The current results are offered to the reader (similar to online bookstores such as Amazon). Some argue quite radically that sophisticated, democratically legitimised rating systems could deprive the “traditional evaluation elites” of their power (Rost 1998a). For journal articles, it is conceivable to combine a rough screening procedure to exclude “libellous, salacious, or otherwise unsuitable material” (Varmus et al. 1999) with a commenting and rating procedure.

The problem with rating is that it is difficult to control for participation. If everyone, including the author, may rate, even more than once, there is a big

chance that we would end up with a strong bias and the result would be highly doubtful, if not unusable. In a closed system with all possibly participating researchers known and with strict access control (everyone having a user-id and password), it would be technically feasible to implement a semi-anonymous system. Such a system would allow checking that each member of the closed group (except the author) has exactly one voice without the individual ratings being de-anonymised. However, it would probably be a major task to administer the groups' membership—a task for scholarly associations. Another problem would be how to cope with controversial results. If a manuscript receives only positive or only negative ratings, a system computing averages would be helpful. But what about a paper that splits the community and gets very positive and very negative scores. The resulting medium score would hide more than it would reveal if the standard deviation (variance) would not be indicated, too. Another solution to this problem would be a more differentiated scoring system, for instance, a multi-dimensional rating scheme that is not only one overall score. Most advocates of scoring systems in academia, however, tend to see it as an add-on to other systems controlling for quality. The same is also true for the next innovation.

(3) *Access and citation counts, use tracking* The quality of a publication can be and is already measured by its use, as well. In the paper world, use can solely be computed by counting sold copies of a journal or book. This is, however, very imprecise, as copies of journals or edited books contain several articles and as many journals go to libraries which have an indefinite number of users. Furthermore, it is indirect, as the sole fact that a copy is within reach is only a very rough indication that the article is actually read and used. Today, counting citations of a particular article in other publications is a de-facto standard and specialist organisations such as ISI compute impact factors. While citations are already a better indication of the actual use of a publication, this method is still not a perfect one as the reason for citing an article can be quite diverse (for an overview on the critique of present scientometrics, see, e.g. Fröhlich 1999; Wouters 1999).

The advent of e-publishing has brought about new perspectives for scientometrics.

- (a) If all publications included in the citation analysis are available online, and perhaps even cross-linked, it is much easier and more efficient to do the citation calculations. In this context, Cameron (1997) proposed a universal, Internet-based, bibliographic and citation database (“citebase”) that would “link every scholarly work ever written—no matter how published—to every work that it cites and every work that cites it”.
- (b) A hypertext environment with standardised meaningful (citation) links would be in an even better position to make the bibliometric enterprise useful (Davenport and Cronin 1990, p. 187 f).
- (c) Unlike in the paper world, there may be automatically generated figures available telling us how many people have already accessed and/or downloaded a particular digital document.
- (d) Finally, use tracking is more than just counting access, it is tracing and computing the degree to which different publications are used and in what ways (Atkinson 2000, p. 67). The data, which are automatically generated

when someone connects to a server with a view to downloading an article, includes more information than only the sole fact of a download, for instance the address of the computer from which the download was requested. This information could be used to establish crude profiles of the users of that document and these profiles could be even more precise in a closed system—for instance in Atkinson's (1996, p. 256) "control zone"—where each user would be identified when logging in. In this online environment, the use of all objects could be tracked and even be differentiated by user group. We would be in a position to answer questions like, which items have been read (or at least retrieved) by which groups of readers (e.g., members of academic departments in the field), which have been read by scholars in other fields, by students, or by the public at large? With a view to indicating something about the quality of a piece, one might envisage even differentiating between the standing of a scholar; if an article is widely read by those at the top of the hierarchy, this might be taken as an indicator for quality.

All of these new tools for scientometrics are not without old (in the case of citation counts) or new (in the case of access counts and use-tracking) problems. To begin with, it still remains dubious whether counting citations or hits does indeed say something important about the quality or performance of a scholar (Remler 2000). Concerning access counts, in particular, there are still a number of technical problems to be solved before access figures would be of any use at all. It is relatively easy to stimulate hits on a Web page or even generate them yourself, mechanically. However, if the practice of measuring Web hits becomes important there will be ways to study adequately patterns of hits that represent real interest, and to extract useful information from these observed patterns (Ullman 1996). Furthermore, not all access is alike. A quick visit to an online article (just looking at the abstract and then deciding that it will not be used) would have to be differentiated from a long look at the full text or a download (Rohe 1998). With certain limitations, this would be technically feasible.

While counting anonymous hits does not involve a danger for privacy, use tracking has such a potential. While the use tracking computers would have to be in a position to match use and identity to make use tracking possible, it would be necessary to prohibit these sensitive data from being disclosed. They should only be presented in anonymised and cumulative form. However, even if this were the case, it might still be possible to find out exactly who read what on an individual level since there may only be very few people of a certain category. While this problem is also present in today's anonymised peer reviewing (the reviewer can often guess who the author is), generalised use tracking may be used to make much better guesses, in particular if data mining techniques are applied.

In conclusion, the digital environment makes it possible, at least in principle, for scientometrics to replace expert judgement in evaluating the merits of a paper. It has been argued that citation ranking and expert judgement of scholars are highly correlated (So 1998). If this scenario becomes reality, the slogan "publish or perish" would be changed to "get hit or get out". While counting papers is certainly no perfect indicator, the Web allows us to measure readership

(Ullman 1996). In the long run, citation ranking and use tracking have the potential to fully replace ex-ante quality control. In the medium run, they might be a useful add-on to current mechanisms.

2.3 Quality labelling and selective databases

Readers of academic publications seek instruments to unburden them from filtering through everything published in the field of enquiry. As we have seen, quality control is one of these instruments. In the paper world, it is the name of the journal or publisher that gives the initiated some indication as to the quality to be expected. If you know the journal, you know what refereeing system is applied. In a bibliographic database, there is no other way of finding out if you do not know the journal. Only if it is a highly selective database, like the S(S)CI, finding a quote in the database is an indication that it belongs to a journal which has been found worthy of being included, at least at some point in time. Many think that this traditional system did not cope with the need for filters in an optimal way. Some argue that the filters were both too rigid (they filter out too much) and elitist. In this view, the present technology is primitive and deficient; therefore, a new technology is needed in order to be able to intelligently select publications (Rost 1998b, p. 8). For the development of an intelligent selection mechanism, the complete digitisation is an important precondition (Rost 1996, p. 174).

Indeed, in an online environment with large archives of full texts we can go well beyond the present state of affairs. Adding labels to each item in a database or archive has two immediate advantages. First, the human reader is in a position to know what control gates each individual item has already passed, and second, this information can be used for semi-automatic retrieval, but also by filters set when retrieving bibliographic entries from the database (Nentwich 1999). The following elements of a labelling system have been proposed.

1. *Related to the quality control system* This includes the simple distinction refereed vs. not refereed; the type and rigour of the quality control system applied; with or without evaluations of appointed referees would be added to the papers; refused vs. accepted (“as a salon des refuses, a working model of what is not currently acceptable, but possibly a latent resource” Davenport and Cronin 1990, p. 187); access figures as a measure for the relative “significance”; and reader ratings.
2. *Content-related* This may include a label for overall importance as assessed by approved editors (e.g. “essential read”—note that this information could be “dynamic” in the sense that descriptions could be added or altered later); type of paper in terms of intended readership or “work level” (Atkinson 1996, p. 257; Ginsparg 1996, p. 6).

Summing up, a digital and networked publishing environment allows the creation of archives and databases that contain not only the bibliographic information and the full text of the published manuscripts, but also indications about the level of quality control it has passed. As a first step, a differentiation between non-reviewed working papers and refereed publications could be made. But there is the potential to go further to indicate the exact amount of quality control applied (editorial screening, editorial review, double blind refereed, open

peer reviewed, ex-post commented and user rated, etc.). The search engines could make use of these additional meta-data with a view to providing filter mechanisms that allow the user to retrieve either all or only those items above a certain level (or items that belong to a certain category).

2.4 Miscellaneous new aspects

Apart from new developments as regards to the procedures of quality control, cyberscience also indirectly impacts the system of quality control as a whole.

ICT has the potential to *Speed* up the review processes. While the referees will still take their usual time (Odlyzko 1994), the overall time for quality check is reduced by electronically distributing manuscripts for reviewing (either by e-mail or via a Web site where the referee accesses the electronic submission in the pre-print archive directly) and electronically processing the opinions and related correspondence (Appel 1996). The OECD report (1998, p. 214) estimates that the turn-around time goes down from the current months to days. A number of projects have developed software to help the editorial process to become fully electronic and faster. It seems almost impossible to go back to the previous slow paper-mail system since we are already accustomed to the new speed.

Due to the increase of non-reviewed publications and the overall increase of available material in the Internet, quality in terms of *Status and Origin* is obfuscated. Layout and appearance may be misleading in the WWW, as it is easy to copy and to make something appear a professional site. Also the Internet address is not always telling, as it is relatively easy to “buy” domain names. Hence, origin and quality are not immediately perceptible. What is then a trusted source? While previously, you had to be in command of a relatively expensive apparatus (a publishing house) to produce scholarly publications, this is no longer the case in the digital age. Academia will have to cope with this by quality labelling (as previously discussed) or trusted source collections which would allow the academic reader to do research in a trusted environment.

The availability of publications in digital format allows for easy “cutting and pasting” which is very convenient for extensive quoting, but also has the inherent increased danger of plagiarism. In particular, this seems to be a growing problem with students writing their home assignment essays or dissertations with the help of official homepages, scholarly online articles or even online collections of essays (Armstrong and Lonsdale 1998). Among academics, some argue that this is probably no real problem; “Cyberspace pirates” (Peters 1996) waiting to copy and pass off as their own an idea for an academic article may not exist. The plagiarism argument has even been turned upside down; the fact that it is much easier to cut and paste from each others works could trigger a new form of dynamic cooperation which could be accepted and appreciated as advancement in communications (Kircz 2001). For sure, this argument is based upon a different picture of science and research than hitherto, namely that of a collective endeavour, solely aiming at the advancement of knowledge, regardless of who is making a contribution—a picture not likely to become dominant in the near future. It is important to note that powerful technical solutions to the plagiarism challenge are possible (Harnad 1998b). There is already software

available that enables us to find out quickly whether large portions of a text are in fact just copied from other Internet sources. In combination with an institutionalised complaints procedure it may also be a way to avoid undue quoting without making a reference. Plagiarism is no inherently technical problem but, as all misconduct, it is a social problem (Kircz 2001).

The e-publishing environment also makes decisions upon *Priority disputes* much easier (Davenport and Cronin 1990) because the exact time of publication can be recorded, at least in principle. As it is possible to fake the date in the digital environment, this would only work in a controlled network with trusted institutions to guarantee its reliability. It has been observed that e-pre-print archives also lead to the strange situation that research groups who worked independently and concluded their research almost, but not quite, at the same time will be credited more differently than hitherto (Merz 1997, p. 260). The winner is clear due to the “time stamp” of the e-archive, whereas the second might not be able to publish the results any more (although this is not plagiarism).

3 The path to a cyberscience certification system

It is obvious that today’s quality criteria for academic work will be of importance for the future e-based publication system. As Raney (1998) puts it, the “message should be clear: No matter the medium, quality cannot be sacrificed for the sake of lesser factors”. The developments described may be considered as a rich tool-kit by which the future system will be built. For sure, there will be huge differences between the various disciplines and fields, depending on working practices and existent quality control traditions (Nentwich 2003, p. 139 ff). The often cited model of physicists is certainly no general model for all fields (OECD 1998). However, the differences in their assessment of quality are not so big that it would not make sense to analyse in general terms the path towards a sustainable certification system.

3.1 Does e-publishing lead to lower quality?

The often heard argument included in the heading comes in various forms and goes as follows: The Internet would enable more people to publish because easy-to-be-attained skills would suffice to put something online (e-self-publishing), whereas previously, researchers needed the help of costly professionals. In particular, an increasing amount of papers that have not yet undergone any process of quality control involving outsiders of the core research group is “published” (in the sense that it is made widely available) via e-pre-print server. This is in contrast to the earlier habit of only “distributing” copies by hand among a few colleagues (which cannot be rated a publication, yet). Additionally, publication would be much faster, but therefore could lead to premature publication of unfinished work. Furthermore, the computer would facilitate “copying and pasting” activities leading to more publications without necessarily more content. All this would necessarily bring about mass production with low quality.

Although there might be some “pull effect” because of the fact that it is easier to publish electronically and to bypass the traditional quality checks, this

argument is not convincing but can be refuted by a number of counter-arguments.

First, “not reviewed” does not equal “no quality” as the prospect of later publication has *anticipation effects*. One can speak of a “self policing” (Harnad 1998a) effect, as the “invisible hand of peer review” (Harnad 1995, p. 5) is also present in the allegedly non-refereed archives since all papers are destined to later appear in journals and hence are written in a quality which might satisfy the future editors and referees. For the physics archive, it is generally acknowledged that peer review operates in an informal way (Odlyzko 1994). But there seem to be safeguards against bad research in other disciplines as well (Varmus et al. 1999; Nathenson 2001).

Second, a high proportion of the increase in scholarly publishing is not due to e-publishing, but a *general phenomenon* with many roots, in particular related to the enormous increase of the number of researchers world-wide. The perception that the digital medium has something to do with it is largely due to the fact that it has improved access to these masses of publications, in particular to the many working papers. However, improved access is at the same time a powerful tool against information overload. Furthermore, we should not forget that even in the pre-cyberscience era, the addressees of working and conference papers knew very well that these were drafts without institutional guarantee that they had undergone some sort of quality control system and had to be treated as such.

Third, multiplication of publications is not always a negative thing and cyber-publishing may be *used in innovative ways*. For instance, writing for different audiences seems perfectly legitimate. More importantly, there is a strong case for at least a partially non-refereed “zone”. Some do not see any danger in producing more, in supporting more discourse, to the contrary: The so-called “publication flood” is less of a problem than ignoring the potential and importance of new ideas—the essence of academia. Probably a large amount of information gets lost in the conventional peer review process, because the end result is only “a single one-time all-or-nothing binary decision” (Ginsparg 1996, p. 6).

Last and most importantly, quality is much *more a question of the quality control systems in place than one of medium*, quantity and speed. Therefore, it depends primarily on the publishing norms in each field (Odlyzko 1994). As soon as an article passed the check, speed of publication or the fact that it adds up to an ever-increasing quantity does not alter anything with regard to the approved quality. There is no difficulty, in principle, to set up quality control systems in an electronic world. The question is: Will they be put in place? And the preliminary answer is simple: Why not? The evidence is abundant that serious e-journals (not to speak of the e-versions of print journals, which have taken over their traditional system anyway) all perform serious quality checks. Many see a natural need for quality control in academia (this is not in the least exemplified by the fears of those who have put forward the initial argument that it might be in danger because of e-publishing). However, there is no reason to believe that academia could not establish quality control in cyberspace. Given the various new opportunities, it is likely that the traditional systems will adapt and develop further. As we shall see, the digital environment will even help to address some of the shortcomings of the present system.

In particular, coping with the increasing output of academia might be done more efficiently in cyberspace than with the bottleneck of the traditional peer-review system.

3.2 Re-establishing and reforming peer-review

While there are some voices arguing that academia could do without quality control altogether, there seems to be a consensus that we should not abolish, but rather reform the present system. Most commentators see no reason why the refereeing system with editorial boards should not be sustained in a digital environment (Davenport and Cronin 1990; Odlyzko 1994; Stichweh 1989; Harnad 1993, 1997). Indeed, there are already many examples in all disciplines of online publications with a strict peer review policy. For instance, the newly established BiomedCentral set up a system whereby editorial boards, whether already established or newly created, review the submission to the archive (Varmus et al. 1999). Note that this is not only something to be considered in the context of publications, but also for scientific databases (OECD 1998). Therefore, the simplest answer to the question how to guarantee quality in cyberscience is that we take over established practices from the traditional paper world and even expand it to new forms of academic publishing.

However, many ask why we should reproduce the paper-based peer-review system in the electronic world. They argue that peer review as practised today is not the right way to go further and should be enhanced and reformed. According to Owen (2000), for instance, Mode-2 knowledge production (Gibbons et al. 1994) and the mixed-mode communication require a more composite and multidimensional approach to quality control than is offered by traditional peer review.

Hence, it is quite likely that traditional refereeing will be transformed and reformed along the path into the age of cyberscience. Some of the new models and opportunities are highly attractive, at least for some groups of actors. Some of the advantages of a networked publishing environment are so convincing that almost everyone will use them. In particular, it seems probable that the tools to speed up the whole quality control process will be universally used. This would mean that, sooner or later, quality control would be a process entirely administered on the Net. From there it is only a small step to actually attach the referees' assessment to the manuscripts, at first perhaps in non-anonymised form. Furthermore, at least for some types of papers (depending on disciplinary field), a combination of a closed peer-review procedure with an ex-ante open peer commentary and/or an ex-post commenting phase seems attractive. This may turn publishing from a unidirectional activity (which starts from the author, passes through a number of steps involving editors and referees, and ends with the readers) into a multi-dimensional communicative process with authors, editors, referees (appointed or self-selected), and readers entering a "multilog", moderated by the editor.

Given the new opportunities of attaching both automatically generated data about the use of a paper, user-generated ratings and quality labels (Sect. 2.3), all of them both readable for the human user and for machines, integrated publication databases will become feasible. These would be highly attractive as one-

stop services with powerful filtering mechanisms enabling the reader to browse and search through different layers of quality. However, this is certainly a big step and not likely to be taken in the near future. There are too many pre-conditions to be met before such a system could be implemented. In particular, e-publishing would have to be generalised, at least in the respective discipline, and it is only conceivable where pre-publication has a tradition. What seems likely, by contrast, is a gradual shift towards semi-automated and more transparent quality control systems.

3.3 Implementation issues

Along the road, a number of implementation issues, including the following, will play a role.

Quality labelling Possibly there would not be any need to make quality labelling in any way obligatory (e.g. via scholarly associations) since there may be other means with the same effect in practice. In particular, social pressure may play a role. Not having a label may be interpreted as a hint that no quality check took place. Publishers might be inclined to apply such a labelling system with a view to promoting their products.

Handling of revisions The new publishing system is likely to allow for subsequent versions of a manuscript. The consequences for quality check need to be addressed. How far-reaching do the changes have to be in order to make it necessary to re-submit it for review? To begin with, it seems reasonable to ask “that each edition must be defined by the author as finished before it can be admitted to the control zone” (Atkinson 1996, p. 259). According to this model, only unaltered originals belong to the core body of the digital library. However, more dynamic models (e.g. “living” hypertext structures) may have a place, too. While I hold that previous versions should remain in the archives, it would be apt for editors to implement reasonable policies when re-reviewing should take place (see e.g. the model of the Living Reviews in Relativity).

Reward system for reviewers Three cyberscience-related developments have led to an increased demand for more and faster reviewers. First, the overall number of publications increased (which is, however, only partly due to the advent of e-publishing). Second, there is the increasing perception that the publication system should be fast and long delays should be avoided (because it is now feasible). Finally, there is the need to make sure that referees take their jobs seriously and that at least some of the self-appointed referees comment in-depth, in particular in an open peer review environment. Some already speak of the advent of a review crisis, even if only indirectly related to the advent of e-publishing (e.g. Owen 2000). It is doubtful whether it would be possible to induce referees to be faster and equally reliable (Odlyzko 1994). A few proposals have been made in this respect. For instance, financial rewards could be envisaged. There are, however, also less expensive proposals. Part of the reward system may be the publication of (revised) referee reports alongside the published article (Harnad 1998a). By this token, something more “tangible” is available for the scholar’s record (to be assessed by promotion committees, etc.). In addition or alternatively, the list of referees could be regularly published on the homepage of journals and publishers. While the option of financial rewards is obviously independent from cyberscience, the latter two proposals

are—although, in principle, feasible in a printed environment—much easier to be implemented in the age of e-publishing.

Policies of scholarly associations, libraries and universities The role of scholarly associations, in regards to implementing novel and reforming traditional systems, is important. For instance, a network of editorial boards and scholarly societies could be made responsible for procedures and criteria (Atkinson 1996). There are some hints that associations have already taken up the issue (Denning 1996; AAAS 1998; Pew Higher Education Roundtable 1998). Others underline the role of the libraries or, more generally, of the academic information services to ensure that national—or preferably international—peer review structures are put in place (Atkinson 1993; Okerson 1991; Lesk 1997). In other words, the involvement of scholarly associations and universities may contribute to acceptance of e-prints.

4 Conclusions

In this article, we have discussed the considerable potential to change the system of quality control, but we have also discovered that a number of hurdles as adjustments of important traditions will be necessary.

Academic quality is, in principle, medium-independent. Whatever the medium, quality may be low or high, depending on the quality control system applied. Cyberscience brings about new forms of quality control, which have the potential to improve the traditional systems, if carefully implemented (in particular as add-ons and only partial replacements of traditional forms). In particular, the new forms of ex-post control are not feasible in the paper-based world of publishing. In some respect, they may revolutionise scientometrics. Working paper archives and journals will probably offer both ex-ante and ex-post open peer commenting. This may turn scholarly publishing into a much more communicative process than hitherto.

Since interactive (as opposed to one-way) communication plays a significant role in most of the new formats of quality control, the single most important factor seems to be time, namely, the time needed for such enhanced communication. It is difficult to predict—and largely independent from cyberscience developments—whether publishing will continue to be ruled by the requirement to publish as much as possible or whether quality becomes more important. In the first case, there will only be very limited or no additional time available for more communication. Hence, the traditional system of quality control would likely move to the Internet, but a qualitative metamorphosis would be less probable. Here, cyberscience's main impact would be that it speeds everything up. In the second case, by contrast, there would be more time available for communication. Here, the potentials could be exploited more fully and cyberscience may indeed contribute to and accelerate this change of scholarly work.

Which of the two scenarios will be actualised, depends largely on the publishing traditions of the various fields and how they develop. The further evolution of quality control systems in the various disciplines would probably not be synchronous. It may well be that in some fields, the more revolutionary

concepts may take shape in the not-so-distant future. In particular, a unified publication archive into which everything to-be-published will already go as an un-refereed pre-print, seems attractive (most of physics has already implemented it, mathematics and the cognitive sciences are following). If combined with quality labelling and if other ratings, as well as comments, were linked to each article in the database, this could turn into a user-friendly publishing system of the future, which would allow for pre-selection of quality levels. In addition, academic literature would be more embedded in the scholarly communication process.

References

- AAAS (1998) Developing practices and standards for electronic publishing in science, a project description. AAAS, ICSU, UNESCO, Scientific Freedom Responsibility and Law Program of the American Association for the Advancement of Science
- Alton-Scheidl R, Schmutzer R, Sint P-P, Tscherteu G (eds) (1997) Voting, rating, annotation–Web4Groups and other projects: approaches and first experiences. Oldenbourg, Vienna
- Appel AW (1996) How to edit a journal by e-mail. *J Scholarly Publ* 27(2):82–99
- Armstrong CJ, Lonsdale R (1998) The publishing of electronic scholarly monographs and textbooks. JISC, Aberystwyth, <http://www.ukoln.ac.uk/dlis/models/studies/elec-pub/elec-pub.htm>
- Atkinson R (1993) Networks, hypertext, and academic information services: some longer-range implications. *Coll Res Libr* 54(3):199–215
- Atkinson R (1996) Library functions, scholarly communication, and the foundation of the digital library: laying claim to the control zone. *Libr Quart* 66(3):239–265
- Atkinson R (2000) A rationale for the redesign of scholarly information exchange. *Libr Resour Tech Ser* 44(2):59–69
- Cameron RD (1997) A universal citation database as a catalyst for reform in scholarly communication. *First Monday* 2(4), http://www.firstmonday.dk/issues/issue2_4/cameron/index.html
- Davenport E, Cronin B (1990) Hypertext and the conduct of science. *J Doc* 46(3):175–192
- Denning PJ (1996) Journals' end? or electronic publishing plan a must. *Comput Res News*, Sept 1996, <http://www.acm.org/pubs/JournalsEnd.html>
- Fröhlich G (1999) Das Messen des leicht Meßbaren. Output-Indikatoren, Impact-Maße: Artefakte der Szientometrie? In: Becker J, Göhring W (eds) *Kommunikation statt Markt–Zu einer alternativen Theorie der Informationsgesellschaft*; GMD Report 61. Gesellschaft für Mathematik und Datenverarbeitung, Sankt Augustin, pp 27–38, <http://www.gmd.de/publications/report/0061/>
- Gibbons M, Limoges C, Nowotny H, et al. (1994) *The new production of knowledge—the dynamics of science and research in contemporary societies*. Sage, London
- Ginsparg P (1996) Winners and losers in the global research village. UNESCO conference, session scientist's view of electronic publishing and issues raised, 19–23 Feb, Paris, <http://xxx.lanl.gov/blurb/pg96unesco.html>
- Harnad S (1993) Implementing peer review on the net: scientific quality control in scholarly electronic journals. International conference on refereed electronic journals: towards a consortium for networked publications—implementing peer review on the net: scientific quality control in scholarly electronic journals 10(01/02) University of Manitoba, Winnipeg, <http://cogsci.soton.ac.uk/~harnad/Papers/Harnad/harnad96.peer.review.html>

- Harnad S (1995) The post-Gutenberg galaxy: how to get there from here. *Times Higher Education Supplement (THES)*, <http://cogsci.soton.ac.uk/~harnad/THES/thes.html>
- Harnad S (1997) Learned inquiry and the net: the role of peer review, peer commentary and copyright. *Antiquity* 71:1042–1048, <http://cogprints.ecs.soton.ac.uk/archive/00001694/00/harnad98.toronto.learnedpub.html>
- Harnad S (1998a) The invisible hand of peer review. *Nature*, <http://www.cogsci.soton.ac.uk/~harnad/nature2.html>
- Harnad S (1998b) On-line journals and financial fire walls. *Nature* 395(9):127–128
- Hibbitts BJ (1996) Last writes? reassessing the law review in the age of cyberspace. *New York University Law Review* 71(June):615–688, <http://www.law.pitt.edu/hibbitts/lastrev.htm>
- Kircz JG (2001) New practices for electronic publishing: how to maintain quality and guarantee integrity. UNESCO-ICSU conference electronic publishing in science 02(20/23), Paris, <http://users.ox.ac.uk/~icsuinfor/kirczppr.htm>
- Kircz JG, Roosendaal HE (1996) Understanding and shaping scientific information transfer electronic publishing in science. Proceedings of the joint ICSU press. UNESCO expert conference, Paris, pp 106–116, <http://www.science.uva.nl/projects/commphys/papers/unescocom.htm>
- LaPorte RE, Marler E, Akazawa S, et al. (1995) The death of biomedical journals. *Brit Med J* 310(May):1387–1390, <http://www.bmj.com/cgi/content/full/310/6991/1387>
- Lesk M (1997) Digital libraries: a unifying or distributing force? Andrew W. Mellon Foundation Conference “Scholarly Communication and Technology” 04(24/25), Emory University, <http://www.arl.org/scomm/scat/lesk.html>
- Merz M (1997) Formen der Internetnutzung in der Wissenschaft. In: Werle R, Lang C (eds) *Modell Internet? Entwicklungsperspektiven neuer Kommunikationsnetze*. Campus, Frankfurt, pp 241–262
- Nathenson C (2001) Electronic publishing and the future of humanities scholarship. *T Int-Z Kulturwissenschaften* 10, <http://www.inst.at/trans/10Nr/nathenson10.htm>
- Nentwich M (1999) The European research papers archive: quality filters in electronic publishing. *J Electron Publ* 5(1), <http://www.press.umich.edu/jep/05-01/nentwich.html>
- Nentwich M (2003) *Cyberscience: research in the age of the internet*. Austrian Academy of Sciences Press, Vienna, <http://hw.oeaw.ac.at/3188-7>
- Odlyzko AM (1994) Tragic loss or good riddance? the impending demise of traditional scholarly journals. *J Univers Comput Sci* 0(0):71–122, http://www.jucs.org/jucs_0_0/tragic_loss_or_good
- OECD (1998) *Science, technology and industry outlook 1998—chapter 7: the global research village: how information and communication technologies affect the science system*. Organisation for Economic Co-operation and Development, Paris
- Okerson A (1991) The electronic journal: what, whence, and when? Presented at the OCLC users council annual meeting in February 1991, *PACS Rev* 2(1):5–24, <http://www.library.yale.edu/~okerson/pacs.html>
- Owen JM (2000) The new dissemination of knowledge: digital libraries and institutional roles in scientific publishing. Workshop “The Economics of Scientific Publication” of the Erasmus Institute for Philosophy and Economics, 19 April, Rotterdam, <http://www2.eur.nl/fw/phil-econ/WESP-jmo.PDF>
- Peters J (1996) The hundred years war started today: an exploration of electronic peer review. *J Electron Publ* 1(6), <http://www.press.umich.edu/jep/works/PeterHundr.html>
- Pew Higher Education Roundtable (1998) To publish and perish. *Policy Perspectives* 7(4), <http://www.arl.org/scomm/pew/pewrept.html>
- Raney RK (1998) One scientist’s view: into a glass darkly. *J Electron Publ* 4(2), <http://www.press.umich.edu/jep/04-02/raney.html>
- Remler A (2000) Lässt sich wissenschaftliche Leistung messen? *Telepolis*, <http://www.heise.de/tp/deutsch/special/auf/8976/1.html>
- Rohe TA (1998) How does electronic publishing affect the scholarly communication process? *J Electron Publ* 3(3), <http://www.press.umich.edu/jep/03-03/rohe.html>

- Rost M (1996) Wissenschaft und Internet: Zunft trifft auf High-Tech. In: Rost M (ed) Die Netz-Revolution–Auf dem Weg in die Weltgesellschaft. Eichborn, Frankfurt am Main, pp 165–179
- Rost M (1998a) Elektronische Foren als Medien wissenschaftlicher Diskurse–Oder: Warum elektronische Foren in der vorliegenden Form nicht für den wissenschaftlichen Diskurs geeignet sind. Telepolis Aktuell, <http://www.heise.de/tp/deutsch/inhalt/on/1389/1.htm>
- Rost M (1998b) Zur Industrialisierung wissenschaftlicher Kommunikation. In: Hartmann F (ed) Informationsgesellschaft–Sozialwissenschaftliche Aspekte, Wien, pp 123–136
- Smith R (1999) Editorial: opening up BMJ peer review. A beginning that should lead to complete transparency. *Brit Med J* 318(Jan):4–5, <http://bmj.com/cgi/content/full/318/7175/4>
- So CYK (1998) Citation ranking versus expert judgment in evaluating communication scholars: effects of research speciality size and individual prominence. *Scientometrics* 41(3):525–333
- Stichweh R (1989) Computer, Kommunikation und Wissenschaft: Telekommunikative Medien und Strukturen der Kommunikation im Wissenschaftssystem. MPIfG Discussion Paper: 89/11, Köln
- Sumner T, Shum SB (1997) From documents to discourse: Shifting conceptions of scholarly publishing. *CHI '98* 04(18/23), Los Angeles, <http://kmi.open.ac.uk/tr/abstracts/kmi-tr-50-abstract.html>
- Tomlins CL (1998) The wave of the present: the printed scholarly journal on the edge of the internet. *J Scholarly Publ* 29(April):133–150
- Ullman JD (1996) Research publication modes need to be reengineered—a discussion. *Comput Res News*, <http://www-db.stanford.edu/~ullman/pub/nopaper.html>
- Varmus H, et al. (1999) PubMed central: an NIH-operated site for electronic distribution of life sciences research reports. Cited 2000 <http://www.nih.gov/about/director/pubmedcentral/pmcprint.htm>
- Wouters PF (1999) The citation culture. PhD thesis, University of Amsterdam, http://www.niwi.knaw.nl/en/nerdi2/group_members/paul/publ/citation/toonplaatje