When it comes to talking about the future of anything, we have to be very careful. It is easy to speculate on the ways in which we might like things to be, or the way things will have to be, but these perspectives often overlook the fact that any way that we think about the future is profoundly influenced by the past, by the ways in which we have come to our existing circumstances, by the accidents of history that have produced the world as it currently stands.

The primary focus of what I want to talk about today is the differences in peer review practice between the arts and the sciences, the difficulties in changing things and the problems of re-thinking peer review within the entrenched systems that constrain our actions. What does a re-jigging of well-known processes do to odious systems of accreditation, such as REF? What might alternative systems look like for the humanities, the area on which I'll be specifically drawing focus for my part of today's talk?

The first thing that it is important to note is that some of the original purposes of peer review have now been lost. This has come about because of the internet and the arise of nonrivalrous commodity exchange (my apologies for using the term "commodity" to refer to your journal articles but, it is, I'm afraid, to some extent the truth). In this new mode, the "ownership" of an item does not come at the expense of another person not owning it. This can be seen clearly in any media: music, books, film etc. It is not longer the case that a perfect, one-to-one copy of an these items can only be made by professionals. While the cassette tape may have played a part in the history, the quality reduction and degradable nature of the medium still made it unwieldy. Now, with a worldwide system of high-speed internet access, exact duplication of these materials is possible and items can be owned without somebody else being unable to retain their copy.

This is, of course, one of the key points that Peter Suber makes in his book on open access and is closely related to many arguments in that field. However, the transition to a mode where there are no longer the constraints of print has wide implications for peer review. Much of the historical function of peer review is the that of the gatekeeper and one of the gatekeeper's functions

was to reduce the quantity of material that was allowed through, not because academics would be overloaded by the quantity of information, but rather because of the page budget in the issue of a journal. With physical commodities, there is a need to limit the intellectual material that goes into the printed artefact because there is a material cost for each page that has to be printed and distributed. This is no longer the case with the internet. Instead, we can distribute more with a negligible increase in cost.

However, this leads to a secondary problem and an evolution in the function of peer review. While many people argue that the function of peer review has always been quality, we can see that, in the past, there was also a monetary aspect ascribed to it. Now, with that additional cost gone – bear in mind, always, that academics are not paid for their review work – the argument has to turn elsewhere. Devoid of economic costs, the focus instead becomes on quality of articles leading to a reduced academic workload. The argument now runs that we must have quality pre-filters so that academics aren't overburdened reading through a mass of undifferentiated rubbish, so that academic journals don't become strewn with nonsense, littered with the cat imagery for which, supposedly, the rest of the internet is famed.

This argument, though, has lead many to speculate whether there aren't, actually, pressing problems with the idea of conventional peer review. Who decides what is "rubbish"? How fair is the current system? What are the intrinsic biases in the process?

To answer a few of those questions, as the answers highlight many of the problems with which the future of peer review will have to contend, first-off, we should ask: is it not objectionable that only two academics, in most instances, although sometimes only one, have the say, in private, with no accountability, as to what is published? It can be the case, especially for early careerists, or those seeking promotion, that their ability to get a job can rest on the private decision of two academics. If one of them is having a bad day and is harsh, it can be the case that the author may be unemployable. Likewise, some papers will get an easier ride based on accidental affinity with the

reviewer's own interests, alignment with their arguments or simply, to be frank with you, upon the mood of the reviewer on the day. Luck plays an enormous part in the fairness of anonymous peer review and it is a system that has us at its mercy.

How anonymous is the system? Theoretically, the review is double-blinded. I should not know the identity of my reviewers and they should not know mine. The benefits of this are clear: the idea is that the review should be impartial and based on the work, rather than ad hominem. Similarly, the reviewers are protected from feeling unable to express their true thoughts in cases where, for instance, the author is a prominent figure in their field. Often, this is utopian. In small fields where people may have presented draft work at conferences, are known for pursuing a specific angle, or through imperfect metadata erasure and/or slips of phrase in self-citation, the identity of the author can be perspicuous to reviewers. It is less often that slips occur the other way around and reviewers largely do keep their anonymity, although I have in the past been able to guess the most likely reviewer of my work.

Furthermore, anonymity has its problems. The shielding of reviewers, with no comeback, can lead them to be harsher than they would be if they recognised that there was a human being on the end of their comments (I've been guilty of this in peer-review feedback before). Furthermore, there is, it seems, something strange about the insistence on anonymity being preserved once a piece has been published. Universities and academia function on genealogies of validation, that is, on hierarchies of prestige and tracing the flow of academic capital through publications. Journals are only as valuable as the genealogies that validate their work as high-quality. This is achieved through submission quantity/quality and rejection rate underpinned by rigorous peer-review. And yet, in the current setup, what we're actually left with is a bizarre situation where, instead of the process of review being visible in order to validate the work, the quality of the review process and the prestige of the people doing the review must be inferred from the perceived post-publication quality of the publication.

In short: under current practice there are only two ways that we know the review was good, both of which are flawed. The first of these is through placing trust in journal brand. While there are strong arguments for why this is a good thing – ie. if the journal is continually publishing good material, then it's probable that their review process is solid – there are also pitfalls here. A "journal" is quite a wide measure of quality over a long period of time. Editors change, journal quality can wax and wane, but our knowledge of this will always be behind-the-times; it takes a while for a drop in quality to register in the general perception of scholars. Furthermore, sometimes prestige can be accumulated simply by one or two articles from junior scholars who then go on to really make it as a leader in their field. Prescience or luck? Who can tell. The second way in which we can crudely measure journal quality – and which surely most affects scholars' perceptions – is by doing our own post-publication review when we read the paper. Clearly, this is inherent in the reading of any academic paper – we use our own academic training to assess the piece's worth and use. We tend to attribute blame (or I do, at least, when I'm not actively thinking about the process) for a poor paper to either the author or to the journal brand. This is interesting; what seems to have specifically failed in this case is the peer-review, getekeeping function. This does not seem to be how it is always perceived though.

So: what is to be done? I'm sure we're going to hear, from the science side, of the innovative practices that have been put in place by BioMed central, so I'll try to avoid stepping on Maria's toes too much, but I suspect that I'll be OK to map out some of the theoretical approaches that have been suggested, while Maria can give us the lowdown on how some of those have worked in the science arena.

Working back up through my proposed chain of problems and beginning with notions of reviewer anonymity, one of the obvious suggestions is to alter the way in which anonymity functions within peer review. To undermine the double-blind process. There are several ways in which this has been proposed, or even implemented. Let us permute through each combination and

consider the advantages or otherwise. Note that I'm still dealing, here, with a pre-publication review process. The first is to remove the anonymity of the author while maintaining the anonymity of the reviewers. This mode seems to add very little. Reviewers would be tempted to judge on the reputation of the person, rather than the merit of the piece alone while still having no comeback upon them for doing so. It also doesn't fix any of the problems and, in fact, is only, perhaps, marginally more honest in the way that it exposes the breakdown of anonymity in the review process as it currently stands.

Conversely, we could remove the anonymity of the reviewers (at various stages in the process) while maintaining the author as an unknown quantity. This mode brings extreme accountability upon the reviewers while protecting the author from pre-judgements against them based on past performance, reputation and stature (at least as much as the current system, anyway). It also gives us the clear genealogy of validation (a term that I first heard used by Martin McQuillan, although I'm not sure if it is one in general currency) present within the extant system and would militate against corruption, to some degree – any links between reviewers and authors would be immediately clear. The disadvantages of this approach are also clear, though. Any system that brings extreme accountability to one side will result in a very conservative environment in which people will be extremely strict in their appraisals, thereby potentially ruling out a whole body of useful work that may be prohibited from seeing the light of day. This may be considered an advantage by scholars who would like a tightening of the review process and to see fewer articles published, but it presents a problem of value judgements to which I will return. Finally, although in some ways this approach helps to deal with corruption in that links between author and reviewers are clear, it might encourage reviewers, in its one-sidedness, to seek the author identity, as there is more burden on them to "make the right call". In short, if they feel its good, but still feel that they're sticking their necks out by reviewing, they might seek verification that it's by a senior professor in the field, a problematic stance. Basically, this tactic exposes reviewers and makes a thankless task

even more risky.

What, then, about making the process completely non-anonymous? Well, there are some advantages here, mostly of the same type as above, but there doesn't seem to be any counterbalance to the elements of entrenched conservativism that could arise from the exposure of reviewers identities. Conversely, they are prone to being overly influenced by the authors' identity here.

So, in each of the de-anonymizing case studies here, during peer review, we encounter problems that seem, to some degree, worse than the original difficulties in the extant model. However, this only applies if we are talking about a gatekeeper model in which these reviewers are deciding whether the paper ever sees the light of day because the journal only wants to be associated, in the first place, with the most exclusive papers, in order to protect its brand. Other journals in the sciences have tried other models. PLOS ONE, for example, has shifted some of their peer review process to a post-publication model. Their peer-review process statement reads as follows:

> "Too often a journal's decision to publish a paper is dominated by what the Editor/s think is interesting and will gain greater readership — both of which are subjective judgments and lead to decisions which are frustrating and delay the publication of your work. PLOS ONE will rigorously peer-review your submissions and publish all papers that are judged to be technically sound. Judgments about the importance of any particular paper are then made after publication by the readership (who are the most qualified to determine what is of interest to them)."

When PLOS was starting out, this caused great consternation among the scientific community, who felt that they would be swamped by a raft of unimportant papers, trivial replication studies and so forth.

It's at this point that I can introduce a not-particularly-subtle plug for a project that I am undertaking, because this issue, derived from PLOS' model, is directly part of our thinking. My project, called the *Open Library of the Humanities* aims to bring a rigorous, solid, digitally-

preserved, low-Article Processing Charge or even APC-free Open Access infrastructure to the humanities. Run with my colleague Dr. Caroline Edwards, we've had support from academics worldwide who have joined our academic steering board, ranging from Peter Suber of the Harvard Open Access project, Kathleen Ftizpatrick, scholarly communication director of the MLA, Martin McQuillan of Kingston, Michael Eisen, the founder of PLOS, Robert Kiley, the Head of Digital Services at the Wellcome Library, John Willinsky of Stanford, Cable Green, the Director of Global Learning at Creative Commons and Vicky Lebeau and Catherine Grant from this institution. The project made the cover article of the *Times Higher* a couple of weeks ago – in which, unfortunately, I was put, through the journalistic quest for a good fight story, into a rather polemic dialogue with David Barnett, who I understand might be here today: if you're here, David, I'd just like to say that I'm sure we're not quite so far away from wanting the same things as that article made out! – has featured in the Chronicle of Higher Education, Library Journal and the London Graduate School site, while we have a piece forthcoming in Research Fortnight. We're also currently in dialogue with major funders to back the undertaking in the region of $1.5m, so it's looking serious.

So, aside from mentioning this because I wanted to get a good reference in and to, obviously, direct you all to our website at www.openlibhums.org, one of the aspects that we're trying to think through is how the PLOS model of "peer review lite" might apply in the humanities. Some aspects of their policy are easy to transfer: "Too often a journal's decision to publish a paper is dominated by what the Editor/s think is interesting and will gain greater readership". Well, I think that the removal of editorial say in what goes for review, beyond a basic check of competency and standard English (or, in fact, whichever language is being submitted – we're thinking through a multi-lingual model and have an internationalisation committee), is probably not that hard to achieve. Other aspects might be harder, though. For instance, some might say that the humanities is founded upon "subjective judgments"; I'm not sure how much that holds, though, because after all, although much of our work is not founded upon empiricism, it nonetheless works on consensus

forming through intersubjective validation, which is akin to attempting to remove subjectivity from non-quantifiable methodologies. I'm aware some might not buy that line of argument, but let's move on. Finally, then, what are we to make of the notion that we could "publish all papers that are judged to be technically sound" while "Judgments about the importance of any particular paper are then made after publication by the readership"? What does it mean for a humanities paper to be "technically sound"? As an initial hypothesis, a "technically sound" paper in the humanities could evince an argument, make reference to the appropriate range of extant scholarly literature, it could be written in good, standard prose of an appropriate register, it could show a good awareness of the field within which it was situated, it could pre-empt criticisms of its own methodology or argument, it would be logically consistent. A paper that did this would be a good start. Secondly, then, I think PLOS are doing something very interesting here with judgements of importance. It may not speak much in their favour that their most popular paper of all time is on bat fellatio, but nonetheless, it is interesting that they have demolished the idea that two people should decide whether a paper is important enough to be published in favour of allowing more material through and then subsequently letting the scientists themselves decide through technological filtering. That seems to me to be great; it avoids the problems of REF-style impact and importance guiding research and instead builds community consensus over what matters, while not suppressing more niche material that may have its own esoteric value.

Slightly differently however, note that PLOS' criteria do not include any reference to novelty. This, I think, might entail some problems in the humanities. A replication study of a humanities paper, beyond empirical humanist research, that is, is contained within reading the work. To replicate the study is to replicate the thought in the mind of the reader and to subject it to disciplinary-shaped critical thinking (with all the problems that entails); ie. an engagement through reading. So, while the OLH project will allow the academics on our boards to decide on this, my preference is to address, as Clay Shirky puts it, the filter failure at the publication side, rather than

casting the vote into the hands of two unaccountable individuals into what sees the light of day, but we will still need to modify PLOS' criteria slightly.

I don't think that it's a coincidence that this shift in thinking about peer review, in multiple places, has arisen in parallel to open access. I think that we are seeing fundamental transformations across all areas of scholarly publication as a result of the internet, but that we rightly fear the damage this may do to an imperfect, but far from dire, system as it stands. Babies and bathwater etc. What I will say is that it is my belief that the shifts we are seeing are not solely the result of wanting to fix problems with academic standards or quantities of work that would overwhelm scholars, but rather a tectonic movement at the economic level, as is the case with open access. Some of the debate around Article Processing Charges in the OA sphere centre around just what exactly publishers do with the vast sums of money that they request. While I share this querying as I think there is a huge degree of inflation here and publisher profit rates seem to sometimes confirm this, a great deal is calculated based on accepted papers compensating for the labour of organising peer review for those papers that didn't make the cut. In short, if you have to filter through 1000 papers for every 1 paper accepted (an extreme rate), the payment for your labour in organising that process of 1001 reviews (note, though: organising, not undertaking, that's still done for free by academics) must be compensated by the APC for the single paper that is accepted. Peer review practice, though, can have a strong impact on these economics. If the rejection rate is cut down by different practices and post-filtering, the logic there is that the cost (per article) should also fall. Somehow, though, I suspect we won't see that, though. Altering practices of peer review are supposedly problematic primarily because of the risk of information overload. I'd like to caution that there is an economic backstory here that might better account for why things are the way they are.

I've only been able to briefly outline some of the problems that are faced in the current world of peer review, but as I've also stated, there are ways of thinking differently about the process. The road to any change is fraught with difficulty – if you allow open review, for instance, especially as it

is a new mode, you can receive very low rates of feedback response (or none at all), as Kathleen

Fitzpatrick warns in *Planned Obsolescence* – but this doesn't mean, wholesale, that change is

impossible. I haven't been able to outline many of the other systems that might work here – thinking

through open commenting, metric-based post-review etc. – but I hope we'll have some time for that

in the discussion. As always, I think we need to be careful that we don't destroy what works in a

quest for mindless reform, but, if there are ways that we can eliminate some of the difficulties, then

we shouldn't be shy about trying them.