The future of scholarly scientific communication

Conference 2015

THE ROYAL SOCIETY

This report describes views expressed at the meeting and it does not necessarily represent the views of the Royal Society. The topics are covered in the order they were discussed during the event.

The future of scholarly scientific communication

20 – 21 April 2015

Day one

Opening remarks by the President of the Royal Society, Sir Paul Nurse

Sir Paul reminded participants that the Royal Society had been in publishing since 1665 when the first Secretary of the Society, Henry Oldenburg, proposed a way of disseminating and verifying new discoveries in science. As a result, *Philosophical Transactions* came into existence: the world's first science journal. This was a truly seminal development whereby scientists from all over the world were able to communicate their ideas, establish priority and, most importantly, expose their work to their peers for assessment. *Philosophical Transactions* established the four fundamental principles (registration, verification, dissemination and archiving) still in use by the almost 30,000 science journals today. But science publishing has remained almost unchanged until a few decades ago and the introduction of the internet.

"The pursuit of science is the pursuit of truth and it is an honourable calling. I don't think we emphasize that enough."

In looking to the future, Sir Paul reviewed some of the key issues the meeting will discuss:

(a) Is peer review 'fit for purpose?' How should it be reformed to serve science as well as possible?

(b) With the widespread criticism of the Journal Impact Factor, how else should we measure scientific quality for the evaluation of individuals and institutions? How can we tackle the intense pressure on young researchers to publish in prestige journals?

(c) How can we ensure that published results are both reliable and reproducible?

(d) Are the mechanisms in place for detecting and dealing with scientific misconduct effective? How can we reform the culture of science to tackle the causes of misconduct?

(e) How might the future look? Data will clearly play a much larger part in the scientific record, but how can we retain dialogue and discussion and how can we ensure that the inspirational aspects of the best published science are maintained? Is society well served by the current models? *On the Origin of Species* was a bestseller, but is modern science writing too dry and technical to grasp the imagination?

(f) What about the 'business of publishing'? How commercial do we want it to be and who (if anyone) should profit?

Session 1A

The future of scholarly scientific communication: what are the issues?

Chair: Professor Geoffrey Boulton FRS

Speakers: Dr Deborah Shorley, Dr Robert Parker, Dr Gary Evoniuk

The Chair emphasized that it was important to be radical in our thinking. There are major changes at play, social and technological and we are witnessing a revolution as significant as Gutenberg's invention of moveable type.

"The ultimate destination of the processes that are now in train is a matter of some uncertainty."

We need to establish what matters and what doesn't matter. What is scientific communication for? What are its essential functions?

Key themes/issues from the speakers:

(a) The enormous cost imposed by publishers on the research community and the large profits made by publishers. Since researchers produce the articles and peer review them, it was proposed that research institutions should take back control of the scholarly communication process from publishers. Learned societies could play a key role.

"Open access is only really tinkering with the existing model."

(b) Many learned society publishers have become very successful and profitable and for most of them, the income from journal publishing is critical to their continued existence. Does their success represent the academy taking back control of its IP?

(c) Industry scientists want to be able to publish all their findings, but they experience great difficulty getting negative results published. Unlike academic scientists, they are less interested in publishing in high status journals and more interested in effective dissemination.

"I really do believe there's a perfect storm brewing in terms of the future of academic publishing."

We should spend more time and effort publishing and mining our datasets than creating text for articles.

"A great deal of data exists already which is often under-utilised."

Key themes/issues from the discussion:

The most frequently voiced concern was the pressure exerted by research assessment processes (especially on young scientists) to publish papers in high impact journals. This is a concern for all scientifically active nations and has created a metrics-obsessed culture which in extreme cases results in cash incentives for publishing in certain journals. The Journal Impact Factor (JIF) has been widely discredited on statistical grounds even as a journal metric and is certainly not an appropriate way to assess individuals. A change in culture is needed so that decisions about tenure and promotion are based on a far broader set of criteria.

"Whenever you create league tables of whatever kind, it drives behaviours that are not ideal for the whole endeavour."

"Why do we do science? It's not to create careers for scientists. It's to increase knowledge for the benefit of mankind. If the need to sustain the careers of young scientists is getting in the way of the primary objective of science there is something wrong in the way in which we organise and motivate those careers."

"The primary motivation of young scientists is to publish in high status journals (whether defined by JIF or something else) and this is a very profound cultural problem."

Scientific communication and scientific publishing are no longer aligned. Funders are very concerned about the lack of publication of negative studies they fund (especially in medicine) and feel the need to change the evaluation process. Publishing data and publishing to communicate should be made to converge – publish the claim and its evidence. It is important to keep in mind the database publishers/creators.

The costs to a research intensive institution of a purely gold open access future would be far higher than the cost of subscriptions.

Do learned society publishers have an inherent conflict of interest in terms of reforming the system? They are seeking to maximise their income from publishing, rather than trying to make the process leaner and cheaper and to institute radical reforms.

"It's difficult to persuade someone of something when their salary depends on them not believing it."

How much is the growth of published articles due to a growth in the science base, and how much is it the huge proliferation of journals? During the 20th century there has been an 'industrialisation of science' as the costs of science and the economic outputs of science have rocketed causing a huge amount of extra financial investment.

"It's more important that the data on which a claim is based are made publicly available than that peer review is maintained. The first is an absolute."

Session 1B

The future of scholarly scientific communication: how might the future develop?

Chair: Professor Geoffrey Boulton FRS

Speakers: Sir Nigel Shadbolt FREng, David Lambert, Dr Ritu Dhand

Key themes issues from the speakers:

Advances in technology and the opportunities offered by the semantic web promise much but we have not really made the strides we had hoped for ten years ago. The 'industrialisation of science', the emergence of citizen science projects, crowdsourcing initiatives and the increasing importance of emerging economies will undoubtedly have a profound influence on the future mechanisms of publishing. The open data movement is also emphasizing the fact that reliability/reproducibility is a hallmark of the scientific process. Unfortunately, real progress is being retarded because so much data is kept dark in the hands of large commercial interests. Where data is open (eg Wikipedia) we can do very interesting large scale analysis projects.

"The publishing process has been shaped by analogue, serial processes. The digital age encourages parallel processes."

There is a prodigious growth in published articles and journal publishers are working hard to adopt new practices to keep the workload of Editors and referees manageable. They are also introducing more data tools into articles and carrying out more checks on statistics and for deliberate manipulation of data and images.

Key themes/issues from the discussion:

Substantial agreement on the need to make better use of technology to build communities to enable discussion between researchers and help in the verification of scientific research. General agreement on the need to better integrate data with publications, but this raised the challenge of how to maintain readability and how this might adversely affect communication of science to the public. Responsibility for data curation, standards and stewardship should rest with scientists, not publishers (as has been the case in molecular biology).

Early career scientists have a crucial role to play as they are 'born digital' and have a more intuitive grasp of the tools now available, especially social media. This will have consequences for peer review which may evolve more towards openness and crowdsourced methods.

Once again, discussion returned to the reward structure and incentives in scientific careers. They need to be aligned with the type of research culture we want to encourage (such as data sharing). Research assessment needs to move away from the narrow focus on papers in high prestige journals and towards a much broader evaluation. The US National Institute of Health (NIH) has adopted the 'biosketch' format of listing a scientist's five or ten most important contributions (not only journal articles).

"Not only are we failing to provide the right incentives, we are actually providing perverse ones."

As long as journal impact factors retain some role in the career development, journals should publish the distribution of their citations. The participants strongly supported the adoption of the San Francisco Declaration on Research Assessment (DORA) by publishers, funders and universities. There was a call for open citation data (rather than having to rely on proprietary sources).

Session 2A

Peering at review: is peer review fit for purpose?

Chair: Dame Wendy Hall FRS

Debaters:
For – Dame Georgina Mace FRS
Against – Dr Richard Smith

Chair: Southampton University pioneered open journals, dSPace and ePrints.org before the web was really present in anyone's mind. We need to think about what the world will be like 20 years or more out. This will depend on what technology is like in the future and we can be sure it won't be like it is now.

Debate 'Peer review is fit for purpose'

FOR: Peer review has the functions of evaluating findings for testability, reproducibility and interpretability. Given the impact on researcher's careers, it must be done fairly. It has served us well over a long period of time; it's not perfect but it's not 'broken.' It uses checks and balances to provide quality assurance and helps authors improve their paper. Journals, in turn, develop a level of credibility and status as a result of how well they do peer review.

Unfortunately, it doesn't always follow this ideal as it is being put under great pressure due to the enormous growth in published volume and also to the sometimes perverse uses to which it is being put. There is however a diversity of detailed practice. Sometimes, things that are blamed on the press are really the fault of how the findings were communicated by the scientist or university press office. There are clear challenges with interdisciplinary papers.

"There are many details to be resolved, but the basic principle of independent, expert peer review seems to be at the absolute core of scientific publishing and I think we need to retain that basic principle, however we go about organising it."

Isn't post-publication peer review problematic as most work is wrong or misleading? The blogosphere is not where most people go for authoritative information. Journals may have wrongly rejected some classic research in the past, but hasn't peer review allowed us to create the pyramid of findings upon which modern science is based? One real problem is that as science has grown, the academies and learned societies have lost some of their reach and influence; they need to reassert themselves in the process to detect and manage misconduct. But none of these are reasons to reject peer review.

AGAINST: Peer review is faith-based (not evidence-based) slow, wasteful, ineffective, largely a lottery, easily abused, prone to bias, doesn't detect fraud and irrelevant. In the age of the internet, we no longer need it. We should publish everything and let the world decide. Peer review in journals only persists because of huge vested interests. There is a dearth of hard evidence as to peer review's benefits. In fact, the evidence points mostly to its detriments. Peer review does not detect errors. A *British Medical Journal* study study with eight errors was sent to 300 reviewers, yet nobody spotted more than five of them (the median was two). The level of inter-reviewer agreement is no better than chance. Peer review is anti-innovatory; there are many examples of ground-breaking work which were rejected when first presented. It's very costly ($1.9bn globally), very slow and time consuming of scientists' time and is needlessly repeated. It is easily abused (for example against competitors) and biased against women and 'low status' institutions. It is poor at detecting fraud (since experimental methods and findings described are usually taken on trust by reviewers).

People aged under 40 have been shown to produce the best reviews. If we are going to retain peer review, it should be wide open (publish reports and discussion). Attempts to improve peer review (by reviewer training and by blinding reviewers) do not seem to have produced any improvement. Career structures are the worst possible justification for peer review.

It's extraordinary that universities and other institutions have effectively outsourced the fundamental process of deciding which of their academics are competent and which are not doing so well.

We should at least subject peer review to some studies and collect some real evidence of effectiveness.

Key themes/issues from the discussion:

Peer review may have value as a 'hazing ritual' by virtue of the fact that the author is prepared to submit to it. Although submitting one's work to the entire world may be seen as far bolder.

It is often claimed that peer review helps authors by improving their articles, although preprint servers seem to do community peer review well and often improve papers in a very constructive way.

There is seen to be a bias in traditional peer review against more original and innovative work which may challenge existing orthodoxy.

Often reviews are not carried out by the eminent expert originally selected, but may be passed to graduate students (though many argue that they do better reviews).

It is impossible to be fair at very high rejection rates and decisions can be arbitrary.

A move to 'objective' peer review was supported as it's easier and more reliable to judge what is correct than what is important/original. However, it can be difficult to get people to review in objective journals as they see it as less prestigious. There is room for both objective and subjective peer review as they achieve different things.

A key issue identified was the lack of good evidence as to the effectiveness of the various forms of peer review and there was a call for more experiments in this area, eg collaborative peer review, post publication peer review with good trackbacks. Most researchers, when asked, believe strongly in the principle of review by peers, but have concerns about its practical implementation.

Abandoning journal peer review completely is a huge step and there is understandable reluctance to do so. Peer review should not be seen as "one thing" – it has changed a great deal over time. We need to really understand why peer review has remained for so long, before we get rid of it.

Session 2B

Peering at review: future developments, evolution and alternatives

Chair: Dame Wendy Hall FRS

Speakers: Dr Richard Sever, Elizabeth Marincola and Jan Velterop

Key themes/issues from the speakers:

A great deal of discussion centred on the need to separate dissemination from verification by use of some system of preprint repository. Physicists pioneered this with arXiv in the 1990s and other similar initiatives have followed, such as bioRxiv. Findings are often, though not invariably, published in a journal thereafter. The key advantages are speed (since material is publicly posted after only brief checks) and the greater level of community feedback and commenting (as there is the opportunity to improve a paper before it is published in a journal). There is also a greater likelihood of errors being picked up (and corrected) due to the much greater exposure pre-publication.

"Decoupling dissemination and certification could solve a lot of problems in the communication process."

Similar experiments are underway at PLOS to allow continuous, open peer review where research contributions are not 'frozen', but evolve, online from the initially submitted version (which need only pass rapid ethical and technical checks, many of which can be automated).

"Much valuable intellectual feedback during the review process is not captured, rewarded or disseminated."

There was a discussion of the high costs imposed by publishers. PubMed Central estimates the costs of technical preparation at approximately $50 per manuscript. Typical hybrid open access Article Processing Charges (APCs) are $3000 which suggests that the cost of peer review is $2950. It was proposed that the peer review process should be done by the scientific community, rather than by publishers. This might be a role for academies and societies.

"Peer review is exceedingly expensive, not peer review itself, but the involvement of publishers."

Key themes/issues from the discussion:

There was strong agreement that a system of universal, free to view and well regulated preprint repositories was desirable. Given how well such systems work in certain communities (such as physics and mathematics), the question was raised as to why journals were still needed in addition.

Once again, the discussion returned to the perverse incentives in the research system in which posting in repositories does not provide credit unless and until articles are published in highly regarded journals.

"arXiv is for communication and journals are for your CV."

Recommendations for improving the preprint repository system were for increased cross searching functionality and better ways to capture all the commenting on articles which may appear elsewhere (much of which is taking place in the 'dark matter' of emails between researchers).

A final point of concern was that market forces, both as applied to publishers and academics actually discourage innovation and experimentation. There was a feeling that many of these crucial issues could be resolved in the Roundtable break out session planned for Day 2.

It was proposed that a more reasonable level of APC should be $300–500 (as is the case with many 'born OA' journals). This is lower than most APCs today (and far lower than per-article subscription pricing). It was noted that the more selective a journal is, the higher the article charge tends to be (as costs have to be recovered for all the rejected articles). One way to offset the costs of rejected papers would be to introduce a submission charge, but most publishers seem reluctant to take this route.


Elizabeth
Marincola
THE ROYAL SOCIETY


Professor Ottoline
Leyser FRS
THE ROYAL SOCIETY


BOSS

The future of scholarly scientific communication

20 – 21 April 2015
Day two

Session 1

What is the best model for assuring quality in science?

Chair: Robert Kiley

Break out groups (mixed by stakeholder category) were asked to consider a number of suggestions for peer review reform tweeted by participants. There was a high degree of agreement between the groups on what a better system might look like.

Most discussion centred on 'preprint' type solutions with some or all of the following key characteristics:

- early deposit of research outputs in repository systems following initial validation checks (many of which might be automated);
- open and collaborative review via commenting or posting of related articles with a view to gradually improving the submission. This might involve social media tools such as voting up or down rather than a binary system of accept/reject. Openness would encourage researchers to learn good reviewing practices;
- credit mechanisms for reviewers (using DOIs to make reviews citeable, awarding of points for each review, ORCID[1] linking of all reviews) which would be formally recognised by institutions and funders as part of a researcher's contribution/achievements;
- data validation algorithms (yet to be developed) would be applied;
- discipline based, operated by international collaborations between learned societies; and
- consolidation/compatibility with existing servers as far as possible.

Such a system would have the advantage of much more rapid dissemination of findings and, if successful, might ultimately render journals obsolete as fewer and fewer authors chose to proceed on to a journal after the repository stage. It would also help to avoid the 'false confidence' placed in articles currently published in journals, as outputs would instead 'evolve' and gradually gain credibility.

Suggestions for approaches based more on the existing journal model included:

- adoption of open peer review as standard (with signing of reviews remaining optional);
- more use of objective peer review (review for correctness rather than impact);
- collaborative review with all reviewers signing a joint report;
- fully portable peer review with authors explicitly 'owning' their reports and able to take them to any journal if rejected, helping to reduce needlessly repeated peer review; and
- adoption of universally recognised systems for crediting peer reviewers with reviewers' scores published to help balance workload. These might be made mandatory by national research assessment systems.

Summary and conclusions

The overall view from the round table was that the principle of 'review by peers' (as distinct from 'peer review' as usually practiced) was necessary and valuable, but should be organised in a different way. In general, participants felt that the opportunities offered by new technology and the web had not yet been fully exploited. There was a role for learned societies and funders to encourage innovation and drive the necessary changes.

1 Open Researcher and Contributor IDs

Session 2A

Measuring science: why do we need to measure research quality?

Chair: Professor Dave Garner

Speakers: Professor Richard Jones FRS, Dr Marion Boland

CHAIR: It's very appealing to rely on the simplicity of a number to assess scientific worth or the value of a journal or the progress of a university.

> Goodhart's Law: "when a measure becomes a target, it ceases to be a good measure."

Metrics are subject to manipulation, so we should look carefully not only at the number, but what it is that number purports to measure.

Key themes/issues from the speakers:

We need research assessment in order to distribute funding between institutions. This is done on the basis of excellence and impact. We use judgements about the quality of research outputs as proxies for the quality of the institution. But research assessment is not a tool for university staff management.

Metrics are useful in informing this process but they should not make the decisions alone.

Most metrics approaches are based on citations, but there is a conflict as the main purpose of research articles is to communicate not to provide proxies for quality. Journal level metrics are not appropriate for this purpose and DORA is to be welcomed.

We should be aware of the disciplinary differences in citation behaviour and the potential for citations to throw up equality and diversity issues.

It is too early to assess the significance and worth of newer metrics, such as altmetrics and article downloads.

Key themes/issues from the discussion:

Assessment of the impact of research can be problematic, especially in pure subjects such as mathematics.

We need to define impact sufficiently broadly that it is meaningful across disciplines and that it doesn't over-emphasize commercial potential of research. We should remember that impact can often arise in unexpected ways.

Scientists need to get better at explaining that our understanding has large gaps in it and therefore pure research has impact merely by adding to knowledge. We should always be able to explain to the public the importance of what we do.

But we also need to get better at defining excellence in ways that do not centre on citations. Or perhaps instead of worrying about definitions of 'excellence' and 'impact' we should think more broadly.

> "What is the deal we do with society that funds us? What is it that we value in the process of research and scholarship?"

Session 2B

Measuring science: how is research best evaluated?

Chair: Professor Dave Garner
Speakers: Dr Steven Hill, Dr Bernd Pulverer, Dr Ewan Birney FRS

Key themes/issues from the speakers:

Research assessment at a national level is not only for distributing funding and for institutional accountability, but also serves to measure the progress of policy interventions.

The biggest challenge is that research is highly complex and multifaceted and cannot be reduced to a small number of metrics. In the UK, the research excellence framework (REF) measures excellence across a broad range of criteria (academic impact and societal impact). Although some parts of government may wish to define impact more narrowly in order to direct research in specific directions.

Although assessment is based on expert review, individual assessment panels are specifically told not to use metrics such journal impact factors. The Higher Education Funding Council for England (HEFCE) has signed DORA. Though of course it's not possible to control the attitudes and perceptions of reviewers on panels.

Good research assessment is expensive and time consuming; we should provide adequate resourcing if it is to be done properly. There is a direct value to institutions in preparing case studies for REF as it helps them to understand better the impact their research has.

The REF approach of using narrative case studies is deliberately broad which tends to mitigate against any temptation to rely too heavily on narrow metrics.

The problems of the (journal impact factor) JIF are well known and are responsible for distorting journal editorial policies and author behaviour (authors are much more likely to cite articles in high impact journals).

> "People game the system at every level and this risks the loss of valuable research in favour of fashionable research."

Thomson Reuters are aware of this and have developed many other metrics, but they have had little traction. Scopus and Google Scholar have helped to break the monopoly of the JIF by providing other citation indices.

DORA attempted to tackle this JIF dominance by recognising that all stakeholders are subject to JIF pressure and should work together to reduce reliance on it, rather than seeking to place blame on only one group. It also emphasises the importance of measuring other research outputs (such as data, mentoring, reagents, software and peer review). It now has 12,000 individual and 500 institutional signatories.

H-index is also problematic, as although it is specifically intended individuals, it seriously disadvantages younger researchers (since it is proportional to the number of articles published). Caution was also expressed about article downloads as a perceived solution to the problem. They are easily gamed and have a low trigger threshold.

We should encourage take up of the NIH biosketch format and we should ban arbitrary JIF cutoffs in institutions.

Molecular biology has shown the way with deposit of structured data which is often without any associated article publication. How can we make this a pan-scientific process?

ORCID has a key role in tracking individuals' deposit, funding and publication. The ecosystem around ORCIDs is developing, it is used by Europe PMC and increasingly by journals. There are now one million ORCIDs issued. It is an excellent way to track an individual's contribution.

Key themes/issues from the discussion:

More needs to be done to drive take-up of ORCID. The Wellcome Trust now make it mandatory for grant applications and the European Bioinformatics Institute make it a part of their HR process for new staff. The REF is also considering making ORCID mandatory.

We need to build a set of metrics that are not citation based (such as data deposit, mentoring students, writing code etc). This will also help to move the focus away from exclusively considering journal articles.

We should forget about ranking journals in any case and focus on ranking articles and individuals. There is no substitute for actually reading articles, rather than relying on metrics.

Attempts have been made to encourage Thomson Reuters to reform the JIF by using median instead of mean of citation counts, but they have so far been unsuccessful.

"Getting away from this obsession with measurement and going back to judgement might be a way forward."

The future of scholarly scientific communication

5 – 6 May 2015
Day 1

Session 1A

Play it again, Sam: why is reproducibility important?

Chair: Professor Alex Halliday FRS

Speakers: Professor Michele Dougherty FRS, Professor Dorothy Bishop FRS

Key themes/issues from the speakers:

In astronomy, the check on reproducibility arises from the application of different techniques addressing the same problem. The importance of open access to data and full description of methods is an essential component in checking reproducibility, particularly instrument calibration

There is an open science project to replicate 100 psychology studies (shortly to be published) which shows that about 50% of studies can be reproduced. Reproducibility has been a concern in medicine for a long time. The responsibility falls on all key players.

Journals: The desire of high impact factor journals to seek 'eye catching' or newsworthy papers leads to a distortion of science at all levels. Funders, in particular were more concerned with novelty rather than replication of previous studies.

Researchers: often statistically inept (either poor methods, or highly complex stats on inadequate data sets), post-hoc hypothesising, p-hacking, poor experimental methods, low fidelity reagents.

Funders: Need to do more. They have been over-emphasizing novel research, not insisting on highly detailed methods in proposals, and often don't require data to be shared.

Employers: still rewarding scientists for publishing in high impact journals, rather than rewarding good science and replications

Key themes/issues from the discussion:

Participants agreed that open data plus a clear description of the methodology and statistics was essential if errors are to be detected. Funders should encourage collaborations in order to avoid generating so many underpowered studies and produce fewer, better ones. Funders should also mandate data sharing to support journals' data sharing policies.

Journals do not emphasize reproducibility enough in the peer review process, they should require better methodology and review statistical methods more rigorously. Many journals still prefer shorter papers, with abbreviated methods and authors may not always want to share their favourite techniques. Journals should also make available the data underlying graphs and tables.

> "People discover things by snooping through a huge dataset and then they apply inferential statistics as it if had been predicted in advance and think p-values are meaningful. They are not."

Most physical scientists use a 5 sigma standard, whereas biologists often only use 2 sigma. Biology journal editors are often reluctant to insist on higher statistical standards for fear of losing authors to other journals.

'Reproducible' does not necessarily mean 'correct' – should we not be focussing more on verification, rather than reproducibility? We also need to ensure that we are comparing the same things (animal strains, cell lines, reagents etc) when making comparisons between sets of findings.

More work is needed to pre-register clinical trials and used standardised reporting protocols (eg CONSORT, EQUATOR).

Authors may be reluctant to share data in case there are errors present. Wider recognition/ acceptance is needed that datasets often contain errors and that exposing this to scrutiny should be seen as positive, not as an embarrassment.

Registered reports is an new journal article format where the decision to accept/reject is made on the experimental protocol and statistics before any data is collected. Not only does this help to address the problem of post-hoc hypothesising and p-hacking, but it also affords the opportunity to get valuable constructive input ahead of the study.

Session 1B

Play it again, Sam: what can be done to improve reproducibility?

Chair: Professor Alex Halliday FRS

Speakers: Iain Hrynaszkiewicz, Dr Matt Cockerill and Dr Nicole Janz

Key themes/issues from the speakers:

Two thirds of PNAS retractions are due to fraud. Better access to original data and research materials and code is needed if we are to effectively tackle reproducibility. We also need more pre-registration, better description of methodology and reporting guidelines and can follow the example of medical research.

Open XML for articles helps text mining and validation. Data sharing is not enough in itself, we need to do more peer reviewing of data and make datasets citeable. Promote CC0 licenses for data. Publishers can influence all these factors.

"The future is already here, it's just not evenly distributed."

Researchers often learn techniques by visiting each others' labs and in scaling up from the lab to industry there is a high failure rate, so we need to address the methodological issues and reduce ambiguity and noise in the system.

We should be capturing data all the way though the process to track the sources of variability in all elements of the method and see what it correlates with. There are lessons to be learned from industrial manufacturing processes in which processes are unambiguously designed up front and sources of error are systematically tracked down and corrected. All the methodology/ process data can then be shared using agreed standards and via recognised hubs allowing others to reproduce the experiment much more closely.

Journals could insist on these standards and develop metrics to reward their use.

"If we want to improve reproducibility we need to look further upstream… biomedical research is more of a blacksmith's shop than an engineered process."

We should do much more to teach students about reproducibility (eg Cambridge Replication Workshop). Motivation for students is that replication is a better way to learn statistics and students' replication studies can often be submitted to a journal as a their first publication. They also learn good research practice.

Universities should encourage replication courses and workshops.

Key themes/issues from the discussion:

The consensus view, often apparent at this meeting, supported positive incentives for scientists to embed the best possible practice in the culture of research. At present there are too many perverse incentives and competitive pressures which lead to unsupportable claims.

Education is clearly of significance and there was strong support for encouraging universities introduce ethics and reproducibility courses. But all players need to accept that a change in culture is necessary if sloppy research and poor practice are to be eradicated.

Capturing more process and methodology data will help make results data more structured and intelligible. This will lead to greater scrutiny of data and identification of errors and reproducibility problems.

Funders and institutions need to ensure that the perverse incentives which have been identified in this meeting are replaced by ones which recognise quality and integrity. There may be a role for governments to show leadership and drive change as they have with open access.

Session 2A

Breaking bad: the scope and nature of scientific fraud

Chair: Professor Ottoline Leyser

Speakers: James Parry, Dr Virginia Barbour and Dr David Vaux

CHAIR: There is a clear connection between reproducibility and fraud in that they are both the result of poor research culture and practice. One of the key issues identified by the Nuffield Council of Bioethics study in 2014 was hyper-competition and too narrow a definition of the success criteria.

Key themes/issues from the speakers:

Research integrity is a problem right across the academic world, but the solutions vary by discipline. A heavy handed approach and focusing too much on major cases can be counter-productive. The more important issue is lower down the severity spectrum than outright fabrication, eg sloppy research practice or honorary authorship. There is also complacency in certain fields of research that there is not a problem and that research integrity initiatives may be seen as box-ticking exercises.

Improving research culture and leadership is a long term game and needs to be by encouragement and standards rather than by a sledgehammer approach to enforcement.

Committee on Publication Ethics (COPE) provides guidelines and advice to member journals, editors and publishers and holds forums where problems are solved collaboratively. They see ethical issues as a core responsibility of editors. We must put in place cultures and processes that enable transparency and reproducibility in the academic literature.

"Academic research is a global industry, but where the norms are set very locally."

There is a disconnect between those who provide the currency of research (journals and editors) and those who measure and use that currency (funders and institutions).

We cannot devolve responsibility just to journal editors, we need to renormalize the professionalism of science. Authors don't commit misconduct to impress journal editors. They do it to advance their careers. The current 'trophy cabinet' approach from funders and institutions means that authors have every incentive to game the process before publication and no incentive to engage after publication.

There are real opportunities for technical advances to be transformative too; publishers should be experimenting with these at the same time as we seek to improve research practice on the ground and develop a shared culture of accountability.

Institutions (many of which are funded by taxpayers) have a responsibility to take this seriously and should publish the number of misconduct cases they have dealt with.

PubPeer and PubMed Commons are valuable online forums for raising concerns about the published literature. The former uses anonymous commenting whereas the latter uses attributed commenting.

Key themes/issues from the discussion:

Journal editors can be reluctant to institute tougher measures for fear of losing authors to less stringent journals. Journals must report cases, not merely reject the article, as this allows the author to take the paper elsewhere and avoid consequences.

Journals also need to provide a better platform for scientific controversy around the articles they publish. Too few journals have a mechanism for follow up debate about articles.

Many ethics committees are not really focusing on research ethics but are instead more interested in protecting their institutions against legal cases. Pressures to commit misconduct often come from principal investigators (PIs) who insist on their teams finding a particular result, no matter what, and offer them rewards for doing so.

Institutions are conflicted and may not always investigate cases sufficiently thoroughly. Journals could help this by publishing the investigations of cases they publish as this will help to expose inadequate investigations.

"'We've spoken to the researcher in question and are satisfied by their reassurances.' – this is not helpful."

Why do universities see this as a conflict of interest? They don't regard investigations of students cheating in exams in this way.

The biggest and most widespread issue is sloppiness and bad practice, rather than deliberate fraud.

Universities should explicitly build in time for their academics to do peer review, work as journal editors and serve on program committees as this will help to emphasize the importance of checking the quality of research.

Session 2B

Breaking bad: the importance of scientific integrity to research and innovation

Chair: Professor Ottoline Leyser FRS
Speaker: Sir Mark Walport FRS

Key themes/issues from the speaker:

There is a spectrum of unreliable science from wickedness through incompetence and sloppiness to unluckiness.

"Because science is a team sport, and increasingly so, there can be a negative incentive for whistleblowers as stigma is attached to the whole team."

A common problem is the underpowered study. The funding system encourages this as it is focused on volume and also fails to encourage repeating the work of others. There is publication bias against negative studies and an even stronger bias against studies which refute earlier work. The obsession of many scientists with the brand of certain journals reinforce these practices.

There is no magic bullet to eliminate misconduct and the policy aim of funders, researchers and institutions is to minimise the incentives which lead to fraudulent behaviour.

There is also a problem with the public communication of science. Investigators can become too enthusiastic about advocating their own work and journalists are often too interested in the most exciting story or the least ambiguous result. The result is that science is often presented to the public and politicians as far too certain and absolute, rather than accurately conveying its inherent uncertainty.

The key priority should be to improve the trustworthiness of science. This is the responsibility of learned societies, universities, funders and to individual scientists. This will involve difficult choices. Should we reduce the volume of research and strengthen the infrastructure? Should we be more rigorous in allowing the brightest minds the freedom to ask the important questions?

We must fund the replication of important findings. We should make use of newer forms of publication such as preprint servers,' beta versions' of articles evolving towards more and more confirmed and reliable later versions.

We must stop regarding it as failure when science graduates don't become PIs; we should stop treating a science education as a vocational qualification. It is a success when scientists go into other walks of life as they take with them the skills of being numerate, analytical and able to communicate. We should recognise this.

"We must focus all of our efforts on ensuring the trustworthiness of the scientific endeavour."

Key themes/issues from the discussion:

Peer review does not in itself ensure correctness. We should recognise that anything we publish is imperfect and the conclusions are uncertain, irrespective of peer review.

There may be lessons to be learned from the humanities where the reviewing process of published works can be and often is highly critical. More open, post-publication peer review would be welcome, but it does not seem to have taken off. It seems that scientists are happy to attack each other anonymously, but far less willing to comment publicly. They may also fear the consequences on another scientist's career or being seen, themselves, as malicious.

Misrepresentations of scientific articles in the media are often not down to exaggeration by journalists, but are present in the original press release from the author or institution. Researchers need to explain the genuine uncertainties in their work and not over-claim. Journals should also provide links to their articles in press releases. The risk of miscommunication is much greater in less formal channels than from professional journalists.

"If we were less invested in our theories, we would be happier to have them criticised."

Calls for funders to insist on higher standards (such as requiring better statistical power in studies) are really calls to the scientific community to scrutinise each others' work more effectively, since the funders use expert panels of scientists to advise their decisions.

Trust in science is not a generic issue. Questions of trust in science arise in a very topic-specific way; usually when science encounters values (eg ethical, moral, religious or economic).

More trustworthy science will cost more as we will have to pay for more verification, managing data and being more open. We will therefore fund less research but may get better value for money as we would be funding more reliable research.

It would help to move away from a 'frozen', 'version of record' approach to publishing and toward a more fluid version of the article which evolves as more information becomes available and further research has a bearing on it. It may also help to encourage better participation in post-publication peer review. But such an approach would require some benchmarking or version control.

There was a call for funding to be made available for experiments and innovations within the research communication space (such as those mentioned during the discussions), perhaps coordinated by learned societies.

Dragon's Den: 'Mending the system'

Host: Roger Highfield

Dragons: Dr Ben Goldacre, Baroness O'Neill, Dr Anita Waard

Pitchers: Dr Cameron Neylon, Dr Kate Hendry, Dr Mike Taylor, Professor Stephen Curry, Dr Aileen Fyfe and Professor David Colquhoun FRS

Ideas for reform of aspects of the scholarly communication system proposed by the pitchers

1) The Royal Society should use its prestige and influence to instigate a cultural shift in how researchers are assessed. It could do so by proposing a change to the system for electing Fellows which would take a broader look at candidates' achievements than is currently the case. Rather than assessing their best 20 published papers, they might consider contributions to a broader range of activities such as refereeing and data curation.

2) An online system should be developed for collaborative peer review using video, audio and text messaging between reviewers. This would lead to a single, open referee review report where no one comment was attributable to any one reviewer (for the reassurance of younger researchers commenting on the work of more senior colleagues). All versions of a paper (including the final one) would be published online. All conversations would be recorded (though not made public) in case of disputes.

3) A search tool should be developed which would allow search across the 4000 institutional repositories and harvesting of materials irrespective of their local architecture. The system would progressively harvest more repositories over time, resulting in a faster search as the system matures, and use open standards to increase transparency. It would also de-duplicate records. An early prototype of this system has already been developed. Estimated maximum cost for full implementation is £4m.

4) Learned societies should collaborate with each other and with institutions and funders to take over all research journal publishing on a fully open access basis with no APCs charged to authors. The money to fund this would come from the government, and be diverted away from universities and funding agencies where it currently pays for journals subscriptions and APCs. Each society would then be better fulfilling its mission to foster communication in the community it serves. Commercial publishers would then focus on other types of publication, such as weekly magazines which might still attract paid subscribers.

5) All research institutions should make a public declaration that they will not use impact factors to assess research or researchers. All journals should publish the citation distributions on which their Impact Factors are based in order to reveal the true nature of the statistic. Thomson Reuters should be obliged to provide the citation data, alternatively it could be provided by Google or crowdsourced by academics.

6) Pre-publication peer review should be abandoned and instead there should be widespread use of pre-print repositories. But the repositories must have adequate trackback and pingback mechanisms to ensure that any comment or discussion about a particular article is fully captured and linked back to the article. Such a system could co-exist with journals (which might still provide an ultimate destination), but ultimately it might render journals obsolete.

Each pitcher received close questioning by the 'Dragons' and (based on an informal poll) the strongest support from the audience was for proposal 3.


Dr Katherine
Hendry

The future of scholarly scientific communication

5 – 6 May 2015
Day 2

Session 1A

The journal article: is the end in sight? Is the journal fit for purpose?

Chair by Professor Sir Michael Brady FRS
Speakers: Dr Aileen Fyfe and Steven Hall

CHAIR: Science communication has undergone considerable change as a result of shifting patterns of power and developments in technology such as the printing press. It would be very surprising if the vast advances in computing and the web did not produce profound changes in the way scientists communicate in the future. The coming generation of scientists were raised with the web and social media and are well placed to take full advantage of it.

Key themes/issues from the speakers:

It is important to realise that journal articles have never been the only source of scientific communication and information. In the past handwritten correspondence and conversations at coffee houses supplemented the scholarly paper. Today, e mail and social media play a similar role.

The journal article became important in an age when printed material was the most effective means of distributing knowledge, but much communication still took place in modes other than articles.

The journal itself might not survive but the peer reviewed paper will continue to play a significant role in the dissemination of science.

Now that print is no longer central, there are far faster ways to disseminate, but are they adequate for the long term preservation of the archive of research? What has not changed is the need for some form of accreditation and evaluation of the published findings. This is currently performed by journals, but there are other ways of achieving the same ends.

However machine readable our articles might become, they will still need to retain some kind of human-readable narrative in order to persuade others of the importance and value of a piece of research.

> "The journal article is evolving empirically, embracing what works and discarding what doesn't."

The journal article is the primary form of formal scientific communication and fulfils Oldenburg's four requirements. It provides a fixed point of trusted (or at least more trusted) reference.

It is evolving to a more fluid form, semantically tagged and linked to many other resources such as datasets and video. It can also be broken down into its component parts according to the reader's needs and is increasingly machine readable.

However, most researchers still prefer simply to read the PDF version and are conservative in their information consumption. They still see the journal article as the primary formal means of communication. But there are many potential audiences and the journal in its current form serves them differently.

Key themes/issues from the discussion:

Perhaps the reason we have not fully realised the potential of technology is the conservatism of the scientific community in their approach to information. This is why the arXiv preprint server seems to be able to co-exist with journals because most physicists want a formal publication after the less formal exchanges that take place on arXiv, as the reward system still requires this. Until we change this system (and publishers have nothing to do with that) the article will remain centrally important.

The journal article must evolve as it no longer reflects the complexity of modern, digital science. We cannot express entire methods and experiments succinctly merely by adding datasets to a written narrative. We must learn to separate content from display and use more dynamic tools, such as wikis, to make the journal article more diffuse and allow it to take different forms according to the requirement of the user.

Collaborative environments like GitHub are not well served by formal publication vehicles. Programmers tend not to 'distil out' articles from their work. Instead, they allow machines to interrogate the codebases for interesting elements. For many of these developers, there is not a formalised career and reward structure that has got in their way. They are not interested in the number of articles they publish.

"If scientific communities become clear about what it is that they want from this new form of communication then we as publishers (or those who will replace us) will rise and provide those services."

The time axis around an article is starting to blur. At the start of the article (with preprint servers) and after the article (with social media and commenting). Perhaps the new additions that publishers are adding are at the wrong point on this timeline and don't fit with the needs of researchers? It is also the case that the PDF version of an article is the lowest common denominator, is standard across all publisher platforms and is compatible with all a researcher's systems.

"Publishing data will become increasingly important and publishing big data is difficult."

The arXiv was introduced (with the backing of the research community) at a time before most publishers were thinking about the web and electronic publishing, but publishers are now starting to work more collaboratively with it and other servers (such as bioRxiv) to build transfer links and connections.

One way of driving change in the reward system would be to reform the academic CV, perhaps by limiting it to a researcher's five key publications, but including other achievements such as lectures, software, fieldwork, data sharing and mentoring. This could be promoted via a DORA-type agreement which institutions and funders could sign up to and might be initiated by a body such as the Royal Society.

Session 1B

The journal article: what might the future look like?

Chair: Professor Sir Michael Brady

Speakers: Mark Hahnel and Dr Anita de Waard

Key themes/issues from the speakers:

We should remember that scientific communication is not only for peers, but for the wider community, especially as science is coming under threat from anti-science forces.

We should be citing claims and data rather than merely articles to provide a better linked knowledge network. We must promote the idea that we need evidence for scientific claims.

> "We live in a world where a tweet from a celebrity can lead to parents not inoculating their children."

Scientists need to remember that they are part of a global endeavour, so they must share their data and do so intelligently. We need to build better tools, software and searching tools to link and integrate datasets and to make them more discoverable. Funders need to recognise the importance of this and fund developments in this area.

Funders increasingly have data policies but they are not providing or funding data centres. This is being left to institutions. It doesn't matter where the data is stored, as long as it conforms to the FAIRport Principles and can be properly linked, discovered and accessed.

The technology already exists, but we need to learn how to use it better. This will only happen if we can reform the system of reward in order to recognise this better handling of data.

Key themes/issues from the discussion:

The possibilities of a genuinely linked web of data are enormous; we will have the ability to check and verify facts in real time against other data sets that exist elsewhere - an 'evidence engine.' An example of this is a system in use during the broadcast speeches of US politicians.

There was a call for every publisher to have a semantic web driven website, rather than just hard-wiring the links between resources.

Natural language processing is very good at 'claim detection' and the identification of trends in data using metadata.

The humanities often implement much more precise and granular forms of citation (eg directly quoting the author) which reflects a difference in the perception of the relationship between researcher and the field in the humanities, as compared to the sciences. The W3C Web Annotation Group is a developing standards to cite multiple anchors to very specific pieces of text or locations.

Many scientists argue, in response to calls to share their datasets, that the data in their particular field is too specific or too idiosyncratic (or simply too large) to be utilised by others and may be difficult to integrate. There is also a distinction between raw data and source data. But the Research Data Alliance exists to find solutions to these sorts of issues.

The research community should insist on open source APIs and tools for data integration. It will also be very important to ensure future readability of data and long term preservation (given the problems we have today in reading data from 20 years ago). There is an initiative underway called the Olive Archive to reproduce many past operating systems on virtual machines in order to allow past software and code to be executable. It is these executable components that are increasingly the product of science.

There was a request that the Royal Society should help to bring these sorts of discussions to a wider audience and foster collaborations.

Session 2A

Scientists or shareholders: Is profit fundamental to a sustainable model? The economics of publishing and sustainability

Chair: Professor Sir John Skehel
Speakers: Mark Thorley, Liz Ferguson and Leighton Chipperfield.

Key themes/issues from the speakers:

For publishers and learned societies, publishing is very profitable. This leads to benefits to society in the form of returns to shareholders and support of charitable activities. But how sustainable is this?

Provided that publishers deliver the services that funders require in a cost effective way, profit is acceptable and indeed helps to fund innovation. But if the supplier has a monopoly (as is the case with journal subscriptions) the market is dysfunctional. Some publishers have exploited this.

Under open access, funders are now buying services and have a choice of supplier. However, as has been raised many times in this meeting, the choice is artificially limited to high impact journals because of the reward system in science. Authors must realise that their choices to publish in high status journals comes at a cost.

"Subscriptions and APCs are part of the total cost of disseminating research"

Publishing services must be delivered in a cost effective way. Publishers should work with funders to develop an effective and transparent market for the services they provide so that researchers and institutions can make real choices about how to spend their scarce resources.

Some learned societies have seen open access as a threat to their income streams and support for their disciplines. We need to support learned societies through the changes that are underway.

Commercial publishers saw an opportunity to enter a market (largely dominated by university presses and learned societies) and bring benefits of scale and organisation to support, manage and finance the process of handling the three million submissions per year. There is a risk in seeking to bring this process entirely back into the academy (with its scarce funding and varying support from government policies).

"Sustainability is an inherent part of being a profit-making organisation."

Many learned societies choose commercial companies as their publishers and they enjoy a responsive and well developed service and see revenues (largely from overseas) to support their activities.

The profit made by commercial publishers has resulted in a great deal of innovation (for example in digital science and open data) and the infrastructure for delivering open access. They have also given rise to large scale collaborations such as CrossRef by providing substantial, upfront capital investment.

Learned societies are charitable operations and use their surplus to fund their objectives. Sustainability is therefore very important to them.

Learned societies collaborate to provide better services to researchers and they benefit from the trust of the communities they serve. They also need to find new ways to diversify and fund their activities as the open access landscape develops. Enabling the dissemination of research is – and should remain – a vital part of the role of the learned societies.

Key themes/issues from the discussion:

Cost of APCs is a concern, particularly the huge range between the various journals and the lack of price sensitivity from authors (as their funders usually reimburse the charges). Hybrid journals often have higher costs than 'born-open access' journals and this is why their APCs may be higher. This is because they may have higher rejection rates (and therefore more cost per published article). Submission fees are a possible solution to this problem, but publishers seem reluctant to introduce them as they feel rejected authors may be reluctant to subsidise the costs of accepted authors.

More generally, concerns were expressed about the cost and affordability of accessing subscription articles and that prices are often being set based on past print holdings. One consequence of this is that learned society publishers often find it difficult to compete with the large commercial publishers who consume the lion's share of library budgets.

There was general agreement that profiting from scientific communication was acceptable in principle, but that value for money must be provided.

Any business not able to adapt to competitive pressures will quickly fail once a better alternative arises. Publishers face the challenge of how to transition from selling product to selling services and how to charge in an appropriate way in order to gain the trust of the community who both supply and consume.

The reward structure in science helps to support the established journals from the traditional publishers as these are generally those with the highest prestige. In China, researchers are provided with cash bonuses for publishing in the highest impact journals.

RCUK prefers 'gold' open access as they see it as a sustainable way forward and they do not wish to set APC limits or restrict the choice of their researchers publishing venues. However they do want to see publishers providing value for money and are concerned about the levels of some APCs.

Session 2B

Scientists or shareholders: the cost to the research community

Chair: Professor Sir John Skehel
Speakers: Dr Stephen Curry and Professor John Wood

In general there was no ideological objection to publishers making profits, but it is a question of how much profit. A number of major companies were mentioned with profit levels between 5% and 20%, but the major commercial publishers had profits over 30%.

This market is not working well as there is not a genuine competition between subscription journals.

Commercial publishers tend to be much more expensive (per page and per citation) than not-for-profit publishers. There are other commercial suppliers to the academic world, but there is more competition between them and they are not benefitting from the free labour of academics.

There are shared interests between academics and publishers (as the former benefit from publishing in journals), but there are also conflicts (such as confidentiality agreements about subscription pricing imposed by publishers on libraries).

Open access is a good way of resolving many of these issues. The two major themes of the meeting are incentives and transparency. Authors should prioritise discoverability of their research rather than publishing in the more expensive, prestige journals. They should be paying for good service rather than for a good impact factor.

The world spends $1 trillion per year on research, but are we getting the best value we can? There are great costs to performing and communicating research, but the latter is getting much cheaper as technology provides new outlets.

Science provides great benefits to human health, industry and society. Costs should not be seen as only to the research community, but to society in general.

The next phase of science will generate even larger research costs. For example the Square Kilometre Array radio telescope will generate 10 times the current internet traffic in terms of data output. The EU is setting up an open science research cloud to help manage the data from European research projects.

Key themes/issues from the discussion:

Support was expressed for the use of preprint servers as a faster and cheaper means of disseminating research.

Funders are reluctant to prevent researchers from publishing in the highest prestige journals (for reasons of cost) as they do not wish to limit academic freedom. In particular, young researchers should not be penalised in this way given the current reward system. But perhaps academics should take responsibility to look for better value for money themselves. Harvard University will not reimburse APCs in hybrid journals. Have their academics suffered as a result?

"We hear a great deal about academic freedom but very little about academic responsibility"

In some cases, seeking cheaper publication routes may be counterproductive as some of those venues may have lower standards and or may be less diligent in checking for misconduct, for example. Many hybrid journals are actually cheaper than some pure OA journals. PNAS maintains a hybrid model as it publishes across a wide range of disciplines and not all have funding available to pay APCs.

Funders, institutions and societies need to take a stand to re-define what constitutes a good journal in which to publish. This should look at a broader range of criteria (such as ethical checking, data policies etc) and not be based on impact factor. There is a role for the Royal Society in this perhaps by creating a DORA-like agreement on what an academic CV should look like.

"The way to develop a cost-effective market is to break the link between publishing in particular journals and measures of esteem of researchers. That's the fundamental problem."

Unfortunately, developing economies appear to be following these perverse reward systems even more intensively as they build their science bases, often wish cash incentives to researchers.

Given the predominance of 'green' open access mandates around the world, we are likely to have a mixed economy of open access and subscriptions for some time to come. It is also important to bear in mind the very large differences between disciplines in how they engage with open access and the publishing venues available to them.

Closing remarks by Professor Sir Michael Brady FRS

There is a complex and rapidly growing ecosystem involving a great many stakeholders in which technology and many other influences are transforming what is possible, what is needed and how it is evolving.

The meeting arose as part of the 350th anniversary celebrations of the Philosophical Transactions which has itself evolved. The Royal Society serves as a forum for debate and as an honest broker on a variety of topics and has convened this meeting in order to be at the heart of these discussions which are of great important to science and to society and which are core to the Royal Society's mission.

Thanks were expressed to all the participants for their time and their contributions.