

Internet Dilettantes' Crowd-Based Peer Review: An Exercise in Mediocrity

Douglas E. Ott, MD, MBA

INTRODUCTION

A suggestion has arisen that posits that a research article peer review process using a crowd-based online protected platform analysis compares favorably with conventional peer review, concluding that it is valid and workable. The suggestion appeared in an observational crowd-based online article filled with guesses, suppositions, and fantasy, justifying a free-for-all peer review. Debate on how alternatives can improve the process, making it more responsive, comprehensive, honest, and ethical and providing an unbiased vetting of scientific reporting would be welcome, but the alternative must meet the standard for scientific rigor. Could be, may be, feels like, and other subjective phrases do not make a change to an article valid. The peer review process gives feedback about misrepresentations, misunderstandings, and the need for clarification. Studies regarding online peer review show that the process has promise, but they have not proved that it is better than the traditional one, nor are they a mandate for change. Self-proclaimed experts with the illusion of possessing knowledge are dangerous for article vetting. Prepublication review should be performed with the utmost attention to detail, the sample size, methodology, statistical errors (such as 1 covariate to 1 response when a multivariate analysis is needed); the use of confidence interval analysis; and validation of the claims made. The reviewer is not the fact finder, but is the fact analyzer and decision maker regarding the text under review. Until there is robust evidence that the current imperfect traditional peer review system will be improved, the search for the holy grail of reviewing processes should continue, but a new approach should not be held to be meaningful if it is not.

School of Engineering, Mercer University, Macon, Georgia, USA.

Disclosures: Dr. Ott is the Chief Medical Officer of LEXION Medical, LLC, St. Paul, Minnesota, and has received a salary for this position since 2000. In this article, there are no references to any product or service LEXION Medical manufactures or sells. There are no references to any product of a company that is in competition with LEXION Medical. Thus, there are no conflicts of interest.

Address correspondence to: Douglas E. Ott, MD, MBA, School of Engineering, Mercer University, Macon, GA, 31210. E-mail: gabiomed@mindspring.com.

DOI: 10.4293/JSLS.2017.00069

© 2017 by JSLS, *Journal of the Society of Laparoendoscopic Surgeons*. Published by the Society of Laparoendoscopic Surgeons, Inc.

Peer review is far from perfect, but nothing is perfect. A recent article suggested using “intelligent crowd reviewing,” differentiating it from “crowdsourced” reviewing as a fix for traditional peer review.¹ How thin can you slice a bad idea? Let me count the ways. Currently there is the standard peer review process and 8 other methods² and the 2 proposed Internet alternatives just mentioned: crowdsourced and intelligent crowd reviewing. The standard peer review method raises some concerns, but it has served dissemination of scientific information well, and the proposed alternatives would lead us into the weeds without a repellent against the crowd, intelligent or not.

THE HISTORY OF PEER REVIEW

Academic refereeing started in the 1750s, when the English Royal Society established a committee to vote on what would be published. To publicize and increase scientific recognition in England, Charles Babbage and William Whewell, around 1830, recommended the commission of reports on all papers sent for publication. With the launching of the *Proceedings of the Royal Society*, teams of 2 or 3 eminent scholars read scientific reports submitted in a process that mimicked the use of expert judges by the French Royal Academy of Science. The French had experience with public judging, whereas the English shunned explicit criticism of a colleague as ungentlemanly. Anonymity won out, with the understanding that it was given ex cathedra (with full authority). During the 1890s the idea that editors and reviewers should ensure the integrity of scientific literature with the referee as a gatekeeper having a duty to science took hold.³ The *British Medical Journal* started giving noneditorial submissions to external experts starting in 1893. The *Journal of the American Medical Association* did not use outside reviewers until after 1940. *The Lancet* did not implement peer review until 1976. In the 1960s, the escalation of public research funding, increased specialization, and competition among journals caused the scientific community's assessment process to morph from a referee system into the peer review process to create consensus.⁴ Peer review, regardless of how it is performed, anonymously or signed, online or by written communication, is fraught with problems of objectivity, bias (recognized or not by the reviewer), reliability, importance, and trustworthiness. The purpose of the peer

review process is to ensure that the work meets standards for quality, that mistakes in methods and logic are eliminated, that results support the conclusions drawn, that no errors are made from citations, that all protocols for human and animal studies are followed and meet institutional review board criteria, and that the work is original and significant.^{4,5} These skills can be taught by a structured methodical training and mentoring process and should be learned by any scientific researcher, practitioner, or reader. Regardless of the method used, peer review takes time and effort and creates bottlenecks, but it is the unifying principle and gold standard for evaluation of scientific material from scholarly publication to grant applications and professional advancement. It is not an Internet open-access consensus statement.

CROWDSOURCED REVIEWING

Diluting the scientific article review process to include an unvetted group of online wannabees, or a responsive group of “experts” leaves more than much to be desired. It has been shown that even the average expert often does no better than guessing, but that “expert” opinion is based more on expertise than on the fickle nature of randomness.⁶ Proven experts make mistakes: they can be biased or taken in or misjudge their own expertise.

Sorting out truth from falsehood is difficult. Reaching a conclusion by groupthink is consensus overriding common sense, and faulty decisions can be made. Aggregation of individual decisions rather than consensus by deliberation is more helpful to the review process. It is the editor's task, by weighing peer reviewers' analyses, to reach a decision about publishing an article. Deciding the merit of a scientific article should not be reduced to reliance on intuition or on cultural or political connections. Separating knowledge from belief is not an easy task. Beliefs that are not based on facts are not knowledge. Information advanced by a scientific article must support and justify its conclusions and must be rigorously separated from opinions. Open minded skepticism, not cynicism; a willingness to accept the conclusion based on trusted data collection; an assessment that reveals new connections to that information in a reproducible manner, with a large enough sample size and power; and following hypothesis-testing tenets and confidence interval assessment are the tests of worthiness for publication. Reviewers must be capable of sorting fact from fiction and fantasy and have a significant base of experience and a wide range of understanding about the subject at hand, not one who claims the mantle without having been

tested. There is a big difference between claiming expertise and demonstrating having actually acquired it. Turning the reviewing process over to self-proclaimed experts as a crowdsourced free-for-all is a recipe for propagation of false and misleading information.

INTELLIGENT CROWD REVIEWING

In a recent expository article, a journal editor calling himself “an experimentalist” reported on intelligent crowd reviewing of 10 manuscripts in parallel with standard peer review, stating that the “conclusion so far is clear: crowd reviewing works.”¹ It is astounding that, with a sample of 10, without scientific rigor, controls, or metrics such a conclusion can be drawn. “The crowd showed at least as much attention to fine details, including supporting information outside the main article, as did those from conventional reviewers.”¹ This special platform of invited referees with various levels of expertise, experience, and specialties were able to comment “on papers they choose to read.” This viewpoint allows the illusion of knowledge and understanding in areas outside the reviewers' expertise, even if they are expert in a closely related subject. It is a matter of not knowing what they do not know. Crowds, even those containing experts, actual or self-proclaimed, overestimate how much they know, claiming understanding, even when it is shallow or nonexistent, creating an illusion of understanding.⁷

It is envisioned that, with intelligent crowd reviewing, “anyone can comment on an openly posted manuscript.”¹ The thinking is that this “anonymous feedback is more candid.”¹ The intelligent crowd forgets that standard peer review is anonymous and candid. The author espousing this idea offers no evidence that standard peer review is less candid than the proposed intelligent crowd review process, saying only that “it works.” A scientist may draw conclusions from a sample of 10, but saying it “works” and is as good as the standard process, does not make it true or valid. The author envisions many expert reviewers congregating on a protected Internet platform. He says that, in the standard method, reviewers are too “overwhelmed” to read and comment on an article and that intelligent crowd reviewing “would lead to faster, more informed editorial decisions.”¹ The real question is why these quickly responsive, willing reviewers are not part of his current standard peer reviewing process. Why is this editor putting up with a panel of reviewers who take months to respond? This is not a peer reviewer problem but a peer reviewer selection problem. Intelligent crowd reviewing has no more likelihood of not being “biased,

inaccurate or otherwise devoid of insight”¹ than does traditional peer review.

Editors and editorial boards of journals have a difficult time in vetting experts, in not overworking the reviewers’ generous donation of time, in committing to evaluate scientific articles using their expertise, and in making a fair judgment about articles submitted to their journal for possible publication. This achievement is no mean feat. Peer reviewing, no matter whether it is called standard or intelligent crowd reviewing, is a process of sorting through proposed material and determining its worthiness for publication, as assessed by experts in the field. There are now too many outlets that will publish anything for a price and the Internet is essentially free for postings. The authors want recognition in a journal with the highest impact factor for their resumé, edification, and future funding. The good news is the free-for-all on the Internet does not have an impact factor. Yet.

Sharing knowledge about experiments, trials, and research and requesting validation by peers is how science advances. It makes the sharing intentional and cumulative and is how knowledge builds on prior knowledge and is sorted out. Not all that is submitted in an article as knowledge is worthy of that mantle, nor is it necessarily methodologically sound or reproducible. Facts, not hyperbole, must withstand scrutiny and rigorous examination. The Internet social media offers little in the way of safeguards and protection against cloning suspect findings, erroneous conclusions, and misrepresentations. The result is scientific gossip, confusion, misjudgment, and prejudice, with calamity as the result. The statistical mean of established known experts recognized and tried by time and history is higher than the mean of the crowd. To be worthy of being peer review referees rather than a pool of unknowns or a group of quick-response “experts” on a secure internal platform may or might not have merit, but with a sample of 10, it is speculative and anecdotal.

DISCUSSION

Editors need a pool of willing, knowledgeable, reliable, dedicated, responsive, functioning peer reviewers. Using an interactive small group by secure text or video conference to exchange points of view and potentially issue a consensus report may have benefit but still requires work on each individual’s part and has no minority point of view in the final analysis, thus generating little more than an “intelligent group review” meeting. Variations of peer review may work, but the jury is still out. The exact vehicle used is not important so long as the end result is

a fair, dependable, defensible vetting of the article in question. An author’s request for a judgment of his or her work depends on and demands an understanding of known science and its application, a respect for inquiry, an objective assessment, and an impartial referee, not a purveyor of personal interest or bias. Other scientists, practitioners of the art in question, and the public depend on a filtering system based on trust and accountability. Knowing the strike zone in baseball, whether a ball is inside the lines in tennis, and what constitutes off-sides in football or performance faults in the Olympics, all require designated referees and judges who know the rules and how to apply them and are known for their assessments, knowledge, and willingness to make the call. Otherwise cheating, chaos, and anarchy prevail, and purpose is lost. Without good judgment, the situation becomes one where anything goes. False, misleading, or inaccurate statements without validity then get perpetuated and difficult to retract. The time to reduce mistakes as much as possible is in protocol evaluation, statistical analysis, ethical conduct, and conclusion validity before publication, not in trying to fix things after publication. Once the genie is out of the bottle, it is difficult or impossible and time consuming to put it back. The whole idea is to keep the pool of collective knowledge as unpolluted as possible. Yes, reputations, careers, funding, and personal egos are in the mix. The advancement of science must not be held hostage to a crescendo of dilution of expertise and falsehoods. The existence of a website or private platform does not ensure the plausibility of an argument. Ignorance and the illusion of knowing create a false narrative that then gets perpetuated. Searching is not learning, seeing words on a screen is not knowledge, scanning or reading a few words is not understanding.

The first peer review congress stated: “There are scarcely any bars to eventual publication. There seems to be no study too fragmented, no hypothesis too trivial, no literature citation too biased or too egotistical, no design too warped, no methodology too bungled, no presentation of results too inaccurate, too obscure, and too contradictory, no analysis too self-serving, no argument too circular, no conclusions too trifling or too unjustified, and no grammar and syntax too offensive for a paper to end up in print.”^{8,9} It is the job of all scientific reviewers to reduce these possibilities to zero. We all want fair, quick, honest reviews on truthful well-designed research. Standard peer review is the worst form of article assessment—except for all others.

Assessors, appraisers, inspectors, examiners, judges, arbiters, umpires, evaluators, reviewers—it doesn’t matter

what you call them—determine that standards are met, rules applied, information gathered appropriately, findings assessed correctly, data interpreted clearly and logically, and appropriate conclusions drawn.

What is desirable is robust science, without dilution or intellectual regression. Every opinion on a subject is not as good as any other. Principled and informed arguments are important to maintain truthfulness. The danger of open Internet peer reviewing is a massive collision of the uninformed, misinformed, and aggressively wrong. A sidebar site such as an open forum may have some appealing attributes, but there is no reason that communication from standard reviewers cannot be just as informative, quick, and open. It takes editorial discipline, leadership, and persistence. Throwing evaluations to the crowd may gather its collective ignorance, biases, prejudices, and narcissistic culture as bases for judging a scientific article.

A small amount of knowledge can make a person think he or she has expertise. Such false thinking leads to advancing of invented or superficial thoughts as “expert” by a person with “expertise” when the conclusions are really stories connected by a false narrative. Expertise should be backed up by facts and fitted to the context of the matter at hand. Without this discipline, the discussion becomes a mirage of loosely supported opinions leading to group-think.

When knowledge is illusory, understanding is shallow. What is required is a causal explanation for the opinion. Causal explanation is shown to temper the illusion of expertise and shallow knowledge and to allow curiosity and new knowledge to be considered.¹⁰ Coupled with the illusion of knowledge is the illusion of comprehension. Superficial understanding and familiarity or shallow cursory knowledge of facts is not comprehension and understanding.

Knowledge is the connection of known facts with new findings from reasoned, thoughtful, scientific observations and experiments. The conversion of something previously unknown to something known through a justified conclusion that can be verified constitutes new knowledge. This is the process that is needed in peer review, whether it is the traditional or digital online type. The medium is not the important factor; it is real expertise thoughtfully applied to the article, giving a fair assessment without bias, and assessing the validity of the conclusions in a methodical manner.

The purpose of the scientific method is to correct, validate, change, and augment understanding of accepted

knowledge and to integrate these observations to expand the comprehension of new facts to a more complete understanding of the subject. The faith in the truths reported in articles is in their replication. An article telling the findings of a yet-to-be-replicated experiment requires ferreting out valid methodology, statistics, and conclusion justification by the reviewer, in a process that is supposed to be the test of scientific truth telling. Real understanding about a subject requires one to be aware of what one does not know and to work to fill that gap.

Not all peer reviewers offer the same level of expert opinions, but over time, editors have been able to judge a reviewer's temperament, fairness, and clarity regarding the elements of scientific study design and assessment. These elements are missing from crowdsourcing. Everyone's opinion is not equal to anyone else's. Some mechanism should decide worthiness for publication. Online publication without expert peer review falls short of that. The Internet is partially made up of persons who think they know better. Experts in specific areas of knowledge have a combination of education, talent, experience, credentials, trust, and peer recognition. Why is there a group waiting on the Internet to pounce on and crowd review an article, but they are not known by a specific editor in an area of knowledge requiring a review? The real trouble appears when the crowd, lacking in expertise, starts to believe that knowing a little bit about something makes them expert or confers expertise. “Knowing things is not the same as understanding them. Comprehension is not the same thing as analysis.”⁷

Crowd reviewing will be a magnet for a massive Dunning-Kruger effect—unskilled and incompetent people espousing authoritative opinions with unfounded confidence. This lack of metacognition creates and fosters a scientific blind spot of not knowing what they don't know, but passing it off as gospel. This is a recipe for the incompetent who do not recognize what they do not know, in that they are less able to recognize their deficiencies and then pretend to possess the necessary knowledge. It becomes a cohort for confirmation bias to congregate. A deep and practiced approach using the scientific method of hypothesis testing and analysis is not one that justifies “evidence” as a finding that an author or reviewer perceives as true.

Opinions of individuals in the crowd are not necessarily knowledgeable. For the individual members of the crowd as a collective, reviewing an article online does not automatically make them peers and capable of critiquing the article in question. “Although the Internet could be mak-

ing all of us smarter, it makes many of us stupider, because it's not just a magnet for the curious. It's a sinkhole for the gullible. It renders everyone an instant expert."⁴ Having facts at one's disposal is not the same as having knowledge and is not a proxy for ability. Are we doomed in the pursuit of peer reviewing online to validate Pommer's law, where the Internet changes a person's mind from no opinion to the wrong opinion?

Let's not have peer review begin after publication. Peer reviewers must be knowledgeable in the subject (depth), skilled in the discipline (experience), cautious (a reasonable amount of doubt), have sustained focus (attention to detail), go beyond system 1 thinking (system 2 thinking), have an active open mind (not closed or biased), stick to their expertise (know what they don't know), possess intellectual humility (know that what they see is not all there is), know when facts change (and change their mind), be able to separate noise from signal (resist wishful thinking), and realize that nothing is certain (two cases are never exactly the same).

The scientific and lay communities do not want misinformation to displace or create false knowledge. The digital electronic public commons is wrought with predictable noise amplified and perpetuated by the Internet crowd. The assumption that a group of random self-identified "expert" reviewers knows more than they actually do is a threat to the well-being of science.¹⁰ There is no guarantee of expertise from an Internet collective review. There are probably fewer errors in judgment from identified tried

and known peer reviewers than from an unknown, untried anonymous crowd.

References:

1. List B. Crowd-based peer review can be good and fast. *Nature*. 2017;54:9.
2. Pater J. Why training and specialization is needed for peer review: a case study of peer review for randomized controlled trials. *BMC Med*. 2014;12:128.
3. Csiszar A. Troubled from the start. *Nature*. 2016;532:306–8.
4. Triaridis S, Kyrgidis A. Peer review and journal impact factor: the two pillars of contemporary medical publishing. *Hippokratia*. 2010;14:5–12.
5. Benos D, Kirk K, Hall J. How to review a paper. *Adv Physiol Educ*. 2003;27:47–52.
6. Tetlock P, Gardner D. Superforecasting: The Art and Science of Prediction. New York: Crown Publishing Group; 2015.
7. Nichols T. The Death of Expertise: The Campaign Against Established Knowledge and Why It Matters. New York: Oxford University Press; 2017.
8. Rennie D. Guarding the guardians: a conference on editorial peer review. *JAMA* 1986;256:2391–2392.
9. Rennie D. Make peer review scientific. *Nature*. 2016;535:31–33.
10. Sloman S, Fernbach P. The Knowledge Illusion. Riverhead Books; 2017.