Soliciting Opinions in Academia: Journals, and Beyond

Alexander Stoimenov

Dongguk University, WISE campus,
Science-Industry Cooperation Group,
123, Dongdae-ro, Gyeongju-si, Gyeongsangbuk-do 38066, Republic of Korea
email: stoimeno@ stoimenov.net
url: http://www.stoimenov.net/stoimeno/homepage

phone: +82-10-4491-4241

Abstract. This article aims to present and discuss a proposed method of evaluating science editors, via authors and referees who have submitted to or (externally) worked for the journal. The main aim of this approach is to gather publicly viewable opinions on editors' exercise of professional duties, with the goal to motivate self-improvement efforts and to provide orientation to prospectively submitting authors. We review first various common problems in journal operation and factors that influence them negatively (such being backlog, and the general lack of viable support for review). We will argue that the proposed scheme of evaluation overcomes, or at least mitigates, these various problems. It offers introducing control on the editorial process, specifically on the quality and integrity of review, in a manner which middles out bias, is protective to those expressing (critical) opinions, and does not interfere with editorial policies. We thus develop a concrete proposal spurring a debate on an effective solution to various problems in journal operation. It aims to contribute to designing a more considerate, accountable, and scientifically sound publishing process. It can be integrated into a broader concept of evaluative fora (an example being for hiring committees), which have arguable advantages of supporting academia compared to both ethics education and legalistic procedures.

Keywords: peer review, academic responsibility, science publishing, teaching evaluations, academic recruitment

1. Need for Oversight of the Editorial process

This exposition is meant as a managemental analysis and not as a statistical survey (unlike, e.g., Huisman and Smits 2017). It aims to motivate an interface in which editors' exercise of responsibilities can be (safely) judged by people they (mostly) work with: authors and referees. Such an oversight process has proved exceedingly difficult to (effectively) establish, yet its lack is the origin of a number of problems in journal operation.

Sadly, these problems are not properly resolvable, and sometimes even exacerbated, by various known factors and recent developments. Here are some most notable ones.

1. Gossip. This universal channel for "funny" stories tends to always benefit most those who can easily get away with it, which are of course seldom the parties with the most merits in a conflict case.

2. The justice system. According to a colleague of mine, the phrase "You can sue me" does seem to have some trading value in the minds of certain editors. How can an author persuade a court that dealing with a particular journal is worthy of filing lawsuits? And how are judges supposed to de-

cide the merits in such a dispute? *Both* systems, the scientific one (Doty 1991) and the legal one (read the U.S. Attorney in Lang 2000, §I.4(b)), are well aware that behavior in the former cannot be properly judged by the laws of the latter. (See further Lang 1993, §V.3.)

3. AI. Among the numerous problematic aspects of this oversweeping trend are paper fabrication (see Skórzewska-Amberg 2022, p.27; Besser 2024 end, Zhang 2024, §3, Case III), and the malfunctioning/mishandling of editorial assistance software (Hosseini and Resnik 2024, Stoimenov 2025). Nevermind, certain people envision "a fully automated publishing process – including the decision to publish" (Horbach and Halffman 2018, p.7).

4. Numerical parameters, like impact factors, acceptance rates, and so on. Apart from creating their own shady businesses (see, e.g., Ibrahim et al. 2025), what they measure is different from responsibility. And they can be easily (and selectively) brought up by editors to suggest tolerating certain (not very proper) practices, like negligence in refereeing.

5. Refereeing as a charitable service. This discords with (a) the, however sad, reality that whenever you apply for jobs or grants, everyone wants to know what you have published – but no one what you have reviewed, as one related issue (b) journals' increasing manuscript inflow, which inevitably pushes down the quality of refereeing, and (c) the enough other people who make (or at least earn) money from the science-publishing process (Jackson 2007, Scarcella 2024). Paid refereeing seen as a moral infringement (Moustafa 2022) or a strike goal (Beecher and Wang 2025) breeds tensions. However, alternatives become increasingly questionable either (see also Zaharie and Seeber 2018). Bringing in free reviews saves editors' costs (and organizing their coverage), but is also employed as a steady presumption to excuse mishaps, with little potential to improve performance. And the demotivation and distrust can only further grow when the system shows little self-sustainability efforts...

6. "Open science" (see Haffer et al. 2019, Wolfram et al. 2020). It appears to promote "open access" also to (presumably) journal-internal proceedings, claiming it as an accountability measure. There is enough evidence, though (see Stoimenov 2025,2026), that such disclosures cannot properly deal with professional or ethical issues (or with both). Allowing public view of submitted manuscripts, referee reports and/or editorial exchanges at best distorts a quality-control mechanism into a self-advertising podium, and at worst bears hazards for people's careers.

7. The attempts (individual, and slightly more organized ones, like Rejecta Mathematica 2009) of unsatisfied authors to go public, against journals' discretion. The minimum problems with this method are that (a) the community tends to easily (i.e., uninspectively) attach a disruptive image onto the "culprits" (see, e.g., the quote of W. Gilbert in Lang 1993, §V.5) and (b) they still risk retaliatory measures from (continually anonymity-secured) referees who see their writings exposed. Similarly unhelpful are authors' direct objections to editors, which can draw angrily complacent reactions (see, e.g., Stoimenov 2023, p.5 top).

8. "If his papers are good, some journal would publish his papers" (Stoimenov 2023, §2, p.5). Widely adopted, this type of reasoning effectively transfers responsibility for operating the process from human beings (who can be held accountable) to ideas that (temporarily) inhabit their minds (which can hardly be done so).

9. Ethics education. It has become recently increasingly common to relate (if not equate) misconduct management in science to teaching courses for students. But efficiency of such programs (Parson 2025), especially for our focus, is riddled with restrictive and predictive variables (see §4). There is more straightforward evidence, conversely, that the community is poorly determined (and com-

ported) to isolate members with questionable believes and behavior (see the quote of Perelman in Stoimenov 2023, §4). And this presents insidious risks that urge real-time solutions.

Each of these counts further raises the question what is a workable regulatory approach. "Workable" should mean features like (a) minimizing personal bias, (b) protecting the ones expressing (critical) opinions, and (c) being taken seriously by those judged as an indicator of what in their work needs improvement.

2. A proposed approach

In pursuit of a solution, I (a professional mathematician) suggested to lean on the concept of teaching evaluations (Linse 2017) which, despite various issues, have stood the test of time. Editors can be evaluated by entering records into a database and displaying their average ratings on individual questions. As well noticeable, many journals themselves present referees to rate various aspects of submissions. A similar procedure appears entirely legitimate to apply in other directions as well.

Such an evaluative setup is warranted as long as we agree that journals have a duty to invest care with submissions they receive. Otherwise, for instance, there seems no reason either to uphold their exclusive handling policy. Of course, journals can counter that they try to return submissions "quickly". But besides that it should be solely on those at the receiving end to decide in how far these policies are effective (which is one of the aspects inquired by the database), a journal cannot speak for others, or for the system as a whole. Hiring committees can also claim that they make most decisions "quickly", and yet they do not bind one to their timeline to apply elsewhere.

2.1 Issues of bias

Extensive critique on editors, at least initially, is a very realistic scenario. And I proposed (Stoimenov 2023) to moderate text comments (if allowed). But think of teaching evaluations. Among others, unpleasant tone of (angry) students is nothing new. Nonetheless, when you go to the administration to justify yourself regarding your poor evaluations, what you will hear is something like "You must learn to (better) get along with students who take your classes." There appears little to contend with a similar principle to be also applicable here.

However, while borrowing from the class evaluation process, we do not try to emulate it. For instance, many universities (properly) deny students' access to final grades (unless and) until an evaluation is submitted. The database has no analogous blockage against authors' result-"attributional" bias (Garcia et al. 2016). But grades would anyway be only one part of a student's semester-long impression about a course. Indeed, Linse (2017, §5.3) reports that students already consciously separate between the quality of process provided and the result given. And with growing experience in the system, (serious) authors gradually move to be less keen on the editorial decision itself than on various criteria how this decision has been reached (*cf.* Horta and Jung 2024, §3, p.4 regarding turnaround time).

Thus, attributional bias is likely at most partially responsible for the evaluation gap in Huisman and Smits (2017, Table 6). I.e., it is reasonably assumable that the review(er)s of rejected papers were perceived as sloppy *not chiefly* because the rejection upset the authors. Note further that the reviewer's anonymity (which has no counterpart in teaching) mostly neutralizes authors' personality leanings. An enhanced way to handle extreme opinions in the database is also to exclude a certain

percentage of lowest and highest ratings from an editor's score average. However, since "extreme" means, by nature, fairly rare, such a cutoff threshold should be rather low (say not more than 2~3%).

Moreover, regardless of their bias, authors need only basic grasp of the publishing process to conclude that an opportunity, like the database, to vent unhappiness serves them much less than the self-damage from presenting weak work. This applies even without submission fees (see end of §2.4) acting as their regulator, since people looking at their submissions are not tight-lipped either (see Stoimenov 2026, §5). In particular, such authors also realize that their opinion will remain isolated unless there are other far more compelling reasons to blame the editors. How many students enroll into a class mainly to dump on the instructor in the evaluation? Such attitude reveals rather soon to the student as utterly unhelpful for their studies, and most students do not see their role at the university in battling their teachers.

2.2 Statistical neutrality

Argumentation is not essential when plurality middles out bias. Students in teaching evaluations (or voters in elections) are not asked rationales either. As well observable, only in exceptional circumstances (like the pre-decision meta review; Stoimenov 2026, §3 end) do journals themselves engage in adjudicating verdicts issued on manuscripts. The common stance (although not always the common sense) is that such investigations would bring the journal's operation to a standstill.

Apart from an equally rightful analogous claim about managing the database, its reliance on subjective opinions is in fact far less exploitable for personal interests (compare Stoimenov 2026, end of §5). While it is obviously (technically) impossible to decide on a submission based on an open-ended community of reviewers, the database may display only average ratings, not individual ones, and every author's verbal comments (if allowed) can be held against anyone else's. This would already provide enough orientation to authors how responsible are editors they intend to work with. (However, since authors pay attention to different aspects of the process, it is better to show averages for each separate question in §5.)

To moderate opinions, one may restrict only one author of a joint submission to evaluate its treatment on behalf of all his/her collaborators. However, once this constraint is placed, the cut (§2.1) of extreme opinions should not include, but preferably break, and even possibly exclude ties [1]. Equal scores will be technically rare, but they may accumulate at both ends of the scale. This must be paid attention to, since such opinions, however downright one-sided, have very independent sources (certainly more so than, say, fellow-students taking a class).

It may be added that for referees, most of these restrictions appear rather uncompelling. Referees are solicited by the journal, with their choice being entirely its responsibility. This largely eliminates many factors underlying manipulatory intent (as long as records are authentic themselves). Positive referee opinions will bear testimony to the journal's stable review process. (Ultimately, a journal is responsible for evaluating what it receives, not for what it receives.) And why would it insist on calling upon people with whom it encounters problems? Likewise, co-reviewers of the same submission may certainly file cases individually, the more that almost always they do not know (about) each other.

2.3 Authors' prerogatives

A related point worth clarifying is that the database should not engage in investigating details like whether/how much the editor indicated by the author indeed handled his/her submission. This is because (a) the database is not purported to meddle in the journal's internal affairs (see §2.4), but at the other end (b) creating submissions mainly for going against specific editors is not a worthwhile bargain for the author (see above §2.1 end), (c) the number of cases per author and journal should be limited (Stoimenov 2023, p.16 top), and (d) exceedingly one-sided opinions can also be handled as explained.

It is noticeable that generally (especially mathematics) journals believe that, whatever the referee did not understand well, it was solely the author's responsibility to have been that much clearer. Such expectation can legitimately be reciprocated. It is very arguably excessive secretiveness (or brusqueness) when the author, who may also have spent weeks or months preparing the manuscript, is withheld until information like who was the handling editor, and what this editor or his/her referee were negative about.

Certainly if the editor specified in the author's database record largely matches his/her subject area, one can reasonably assume that this editor at least had (or should have acquired) internal access to the submission's progress. In addition, manuscript transfers should not be misused for editors' evasions. (Journals would hardly care either if an author tries to divert critique referring to coauthors.) Furthermore, since an author's highly subjective views are little persuasive, it suffices to uphold the case-count limitation (c) only to the filing author, if he/she acts on behalf of all coauthors of a joint submitted paper (*cf.* §2.2).

This can also be related to §1, 5. It is hard to dispute many editors' assertions that "good" reviewers are difficult to find. But seeking ways to raise the work's appeal for broader commitment is long-term likely a more productive strategy: a good editor values a diverse reviewer portfolio, preventing a handful of individuals from stamping their taste onto what the journal publishes. In addition, editors like stressing *their* view of what a "good" reviewer means. With the opinion of both authors and co-reviewers effectively unavailable, this can seek verbal justification of subjective choices which are not always optimal, and sometimes overtly inappropriate.

Authors *suggesting* reviewers supports obvious bias (see Skórzewska-Amberg 2022 for extreme misuses), and – explicably – remains an option at very few journals. But *opposing* reviewers more legitimately merits assessment. In certain situations, the editor knows about a reviewer's conflictful relationship with the author – though own past experiences, or even the author's explicit indications (see the monograph case in §2.4 below, or Stoimenov 2022, p.594 bottom). Disregardfully, he/she can keep on asking this same person – partly for convenience's sake, and partly because "it is hard to prove it" (Stoimenov 2022, p.590 l.8). What should one then make of this editor's rhetoric in defense of reviewers? There are countless unfavorable outcomes across society of "good" people being shielded from criticisms -- including what has given rise to review itself.

Editors should, therefore, seek first at themselves the origin of criticism they face. It is intended as a prompting to take their duties more seriously, not as a personal offense to set straight.

Satisfied authors have a reason to vote either: they credit the service they received, and suggest other authors to consider using it as well. Reviewers' opinions will achieve further balancing: the fact they are both external and have little at stake regarding a submission lends a high degree of authenticity to information they provide. When performance is deemed appropriate, then neither student ques-

tionnaires nor internet rating fora have proved to generically degrade into libelous squabble rooms. Actually, commendable editors can acquire this way evidence for their credentials.

2.4 Independence of journals' operation

A further relevant point is that, regardless of opinions expressed, filing cases does not overturn editorial decisions. This is similar to teaching evaluations and assigning final grades being separated processes, and also keeps distance from certain authors' individual contesting activities (see §1, 7). Thus, the burden of proof how much his/her work is worth remains no less on the author. According with this, at least to large extent justifiable, expectation of journals, the database further disavows their potential complaints about such method of gathering opinions about them.

It must be emphasized that this expectation is *largely* justifiable – simply because journals do not restrict who may send what to them -- but by no means white-clean. Its excessive defense may mount narratives like §1, 8. Another telling caveat was the monograph case of Stoimenov (2022). It shows an editor unceremoniously following dismissive advice from people when the author is trying to correct their published flaw. Among these people was, in fact, the one who had been suggesting to this same author the editorial philosophy just referred to.

When a disputable rejection occurs at journal J, the author X has no effective means to convince the editors, except by eventually getting the paper published. And while this has happened not once or twice (and often enough in some more or less comparable journal), it usually follows a further indefinite sequence of (review, etc.) delays. Until then, the editors of J can safely assume that their arguing is long forgotten. And more generally, X would do better looking ahead, rather than dwelling on (dreary) past events. But even if X then, nonetheless, informs the editors of J about the publication, the reaction is less than sympathetic (Stoimenov 2023, §2, p.5; compare with the other pontificating Stoimenov 2022, p.590, l.6-7).

Of course, during his/her extended ordeal, the author has the continuous opportunity to improve the manuscript. And no one can or will probe whether, when it was accepted by one journal at one time, it was worth rejecting a previous version by another journal at another time. The Solomonic shorthand of leaving the manuscript's publication entirely the author's responsibility may be inevitable. But one must also admit that it naturally provides another source for editors to enjoy (and if that does not work for some reasons -- enforce) their vindication for questionable gestures -- especially over corrections [2]. And the database's incapacity to arbitrate such developments either may, on balance, still effectively concede more to editors than to authors.

Review brings its own unknowns. Fundamentally, like editors, referees want to see quality, and there is little to shake about that attitude. But demand for quality can too easily mix with (and become a mask for) other concerns. Particularly in mathematics, arguments are notoriously difficult to follow, and the lack of proper support for review (§1, 5) further fuels the unexcitement. This has, for many referees, put up front personal preferences which, also beyond mathematics, are provably close-to-random (Loewe 2022).

Referring to taste explicitly would be a too revealingly unsolid argument, thus other formulations are employed, among which the family of verbs "seem/look/appear" is archetypical. In the function of selling statements without (effort at) verification or argumentation, they will leave in research a ref-

eree little impressed. Yet the editor treats this same dictionary used in a review as something normal [3].

The gap between what submissions are exacted to deliver, and what reviews are condoned to afford, can be self-reinforcing. And since nothing balances how often one takes either side, this can breed tension (see Stoimenov 2026, §5). Of course, the stakes are different as well. But to see the role of the database, consider again the analogy to teaching. It is only students who can receive credits for a course. This does not mean, though, that the instructor, or a TA, can treat them as he/she pleases, and class evaluations can rather effectively expose such arbitrariness.

Manuscript processing which, at least in the author's perception, is often enough undeserving (or even unlucky) also makes it plausible – and, worse, perhaps even advisable (see also Stoimenov 2023, §3.4 top) – that authors do not take very seriously journals' talk over rejections. This only adds reasons for a "shoot-until-it-sticks" mentality. For a number of materials I was shown for review, I can reasonably assume that the author would not have sent them in, at least in that form, had there been some mechanism for more self-attentiveness.

Submission fees (see Stoimenov 2026, §4) could act as such mechanism -- including against "automated" (see §1, 3) receptions. However, it is only through some provision like the database that they will help fostering discipline on either side. With extra revenue (in addition to other author payment options) but no extra control, journals' self-intendance would improve moderately at best, and authors will likely start abandoning submission-fee-levying journals.

3. Channels for opinions

The formulator of the editorial paradigm of §1, 8 tried to argue about it by referring to "[o]ne world famous mathematician in our department" making an analogous (evasive) statement regarding a job candidate (*cf.* Stoimenov 2022, p.592 bottom). Such situations suggest that the outlined concept should not be restricted to publishing.

Allowing applicants to evaluate hiring committees appears somewhat impractical: an applicant is unveiled, in average, too little about a hiring process to be asked vigorous assessment. (At least, to his/her advantage – and unlike submitting a manuscript – he/she can apply to several places simultaneously.) But every member of a hiring committee is in a well-suited position to judge every other one. Through an evaluative provision, faculty in charge would be able to express (anonymously, and hence) non-repulsively their outtake on a question like how far (they think) did someone such as this "world famous mathematician" help a fair recruitment process. Some sample criteria are: "Did the member study the applicants' background and assets properly?" "Did he/she try to understand the applicant's situation well during an interview?" "Did the member soundly argument his/her judgments?" "Did he/she accord well with other members' opinions?" (For appointment references, see Stoimenov 2025.)

Of course, there are issues to consider, e.g., who and how should submit and view such ratings (to gain helpful insight from them)? Editors are meant to be widely approachable, but panel members less so. Similarly to the journal database, some management would have to legitimize access based

on proof of employment and committee co-participation. These seem no unsolvable problems, however, we cannot accommodate their extensive technical discourse here as well.

Still, the caution is relevant that a feedback-gathering structure, for whatever sort of academic activity, should *not* be *university-internal.* I suggested (Stoimenov 2025) class ratings to be managed by some body like the student union. This can be regarded as "internal", however, students are very little of the policy- (and politics-)makers at their place. Responsibles at an institution cannot be relied on to properly handle critique internally. Pressured with income and (certainly local) careers on the line, they have sometimes, however unintendedly, but effectively, contributed to ostracism and cover-ups (see, e.g., Lang 1993, §I.2, §V.2 end, and the quote of W. Gilbert in *ibid.,* §V.5, and compare with Nature 2018, 4th paragraph).

The point behind suggesting these further applications is that creating such channels for opinions may, hopefully, draw (or get drawn) certain scientists away from considering *how little trouble* (they reckon that) they will have with what they do as their self-approval criterion. This level of stimuli are particularly detrimental in science (see Doty 1991). Thus, the intended purpose of the proposed evaluative fora is *to support science community manage itself better.* As such, they promise some advantages over alternative (sometimes infamously) well-known approaches.

4. Courts, courses and close-ups

Since the 1980's, where major misconduct cases started surfacing in science, the understanding was promoted (and cemented) that venues to settle behavioral disputes be organized in legal-type formats (*cf.* §I, 2). This is witnessed in diverse coverage about "trials" (Nature 1991), where "accusers" seek to "punish" the "accused" (see quote of Kevles in Lang 2000, Footnote 4, and Culliton 1991), and lawyers mediate incl. in the function of "judges" (quote of Kevles in Daemmrich 2014, "Investigation #3"). Such people (apart from, for instance, offering quality of service roughly proportionate to the wallet of the client) are naturally trained within a civil law background.

Attempts to police science within this setting get marred with contradictions (see Lang *ibid.*, Footnote 2 as an example) and absurdities (like the approval of practices as soon as sufficiently many scientists follow them; *ibid.*, §I.3(b)), and do not correlate well with efforts like maintenance of published literature (see Daemmrich *ibid.,* i.p. Endnote 16). Plaintiffs are expected to present severe evidence. And if such is inaccessible, or if defendants get exonerated, everyone can be prodded to pretend as if nothing wrong on their part occurred (and occurs). This perfidy of legal(istic) challenges puts a particularly stark contrast to how behavior in science should be evaluated (and regulated).

The recurring notoriety of such face-offs is fittingly characterized by one of those (feeling) affected (O'Toole, addressing U.S. Congress, as quoted by Weiss 1989, end): "This is the wrong forum. It is just that we are stuck here because we couldn't work out a better one." Compare with the preceding hardship of Stewart-Feder (Lang 1993, §I.3), and see de Russy and Langbert (2005) for further cases. In the monograph incident (Stoimenov 2022), S even offered a partial remedy of the error he found, but met the authors' adversity (which continues to this day).

Instead, the present proposal better aligns with "the court of public opinion" (the U.S. Attorney, as quoted in Lang 2000*,* §I.4(b)). It may more effectively tackle some types of problems in academia (not universally, but clearly far beyond only journals [4]) as long as they are small. The alternatives of turning cover-up into one's profession's craft, or boosting confrontation with opponents (*cf.* Lang 1993, Conclusion), neither brings it.

At this occasion (and as a digression), it may be emphasized that ethics education is another rather ineffective measure. Suspicious notabilities left to do well (see §I, 9) can all too easily have – and in certain cases are directly given (O'Toole 1991, p.183 right) – opportunity to "educate" their views. "Who would determine the content of courses in scientific ethics [and w]ith what choice of material?" (Lang 1993, §V.4(b)).

Teaching can leave students satisfied (Parson 2025) even if withholding relevant insights (Lang 2000, §II.3), or offering "carefully wrought" (Lang *ibid.*, §II.2) yet externally disputable ideas (or ideologies). A student preparing for such a course could read, for instance, a bookful on an actuality like Baltimore's and still learn nothing about "the obligations of scientists when questions about their work arise" (Gunsalus 1999). Should this student then "truly believe that science benefits when a prominent scientist acts as Baltimore did in this case?" (Gunsalus *ibid.*) The road to hell is sadly also paved with "good intentions" (Lang 1993, §V.4(b), quoting J. Swazey).

Sklamberg's valuable essay (1990) presents a student's first-hand testimony documenting the confluence of pedagogic and demagogic in political science. How do we ascertain that the ones advising strive for higher standards than those advised? Zhang (2024, §3, Case II) discusses a recent example, which also (dismally) corroborates what Lang (1993, §V.4(b)) wrote more than 30 years prior regarding the "mentor system".

Besides, when measuring effects of such processes, students later involved in academia (and in its) management form a rather minor, and hence not reliably a statistically indicable, subset. This group is hard to encompass in advance, as are the variety of future individual experiences and circumstances that will influence their professional behavior. Despite the reiterated emphasis lent to the topic (de Russy and Langbert 2005, Parson 2025, etc.), it transpires that teaching promises very suboptimal ethics-enforcement solutions, and why more practical directives (see, e.g., Nature 2018) follow a different path.

On another related venue, of course, there are similar surveys conducted by publishers. They are inefficient for our goals for a number of reasons. First, they partly concern issues, like printing quality, which are (relevant to the publisher but) not directly linked to the editorial process. Next, they are often restricted to authors of successful submissions, they usually do not address editors personally, and neither are results disclosed. Moreover, in general, recruiting or replacing editors is a strenuous task for a publisher. This means that he cannot be assumed entirely impartial regarding his editors, and therefore is not the ideal party to manage information about them for public use. (He is very much, though, encouraged to do so internally, and lift some of the aforementioned limitations; see Stoimenov 2023, §5.1.)

Of course, emphasizing on the human factor can be criticized. But, for instance, the same manuscript submitted at the same time to the same journal may also elicit different decisions by different editorial boards (or even under different order of receiving reviews; see Bornmann and Daniel 2009). Despite such inherent ambiguity, the authors/readers of the journal can still be satisfied when the editors (whichever ones) make a genuine effort to maintain behavioral/scientific standards. While the development of AI remains to be followed, at this point, I can imagine certainly a number of mathematicians plain dismayed with the style of automation of §1, 3: (accelerated) determinism is not imperatively a formidable ethical or quality-evaluative instrument.

5. The Questionnaire

The aim of this note is to spur a debate on what, specifically in mathematics, would be a proper model and how to implement it, and to present this concrete proposal. The questionnaire for evaluating editors is given in Stoimenov (2026, §9), but is of course largely up for discussion.

In order such a project to be effective, it needs community support. People who do not, or have not done, their job properly will regularly try to cast doubt on a source of complaints, particularly by emphasizing on its deficiencies. The key question about their exercise of duties gets often blurred in, sometimes deliberate, controversy, especially when outreaching professional circuits. Journalistic standards, like the quotes in Lang (2000, Appendix I), can be left to the public. But even prestigious academia appraising an account as "explosive" rather than accurate (Lang *ibid.*, §II.2) does warrant some observance. To prevent such turn of events, a regulatory activity must be run by a group of respected mathematicians, and must be launched from the start with a very clear-cut concept.

My articles (2023,2025,2026) contain further details on how such a process could function. For instance, the question to referees about co-referees pertains to (the journal providing opportunity for) mutual evaluation. In fact, the questions to the author about the referee can very much (and likely even more pertinently) be posed to any co-referee. But this level of detail may be more appropriate for a journal's internal inquiry (to improve its future choices whom to consult). Thus, the database intends to limit itself to the given one general related query. The question how to report and probe "misconduct" is also extensively treated in Stoimenov (2026).

Notes

[1] One instance where the cut is decided by a number of people, but those missing the cut exclude ties with scores within that quota, is at golf tournaments.

[2] A common pretense (see Stoimenov 2023, §1.4) is that such material is unsuited for refereeing. Baltimore's remarks on Stewart-Feder (as quoted by Lang 1993, §I.3) show that this is exploitable to debase scientific objections.

[3] As documents Stoimenov (2022, p.589, l.17), certain people contemplate using this tactic *publicly*, effectively to spread baseless havoc regarding the validity of publications.

[4] The mentioned cases suggest, for instance, that similar to class evaluations can be done with student project/thesis supervision.

References

Beecher, K. and Wang, J. (2025), "Peer referee fatigue, or peer referee refusal?", *Accountability in Research* **32(5)**, pp. 838–844, https://doi.org/10.1080/08989621.2025.2463977

Besser, L. (2024), "Wiley's 'Fake Science' Scandal Is Just the Latest Chapter in a Broader Crisis of Trust Universities Must Address", online article, ABC News (Australia), 20 May 2024, https://www.abc.net.au/news/2024-05-21/wiley-hindawi-articles-scandal-broader-crisis-trust-universities/103868662?utm_source=abc_news_app&utm_medium=content_shared&utm_campaign=abc_news_app&utm_content=other

Bornmann, L. and Daniel, H.D. (2009), “The luck of the referee draw: the effect of exchanging reviews”, *Learned Publishing* **22(2)**, pp. 117-125. https://doi.org/10.1087/2009207

Culliton, B. (1991), “NIH need clear definition of fraud”, Editorial, *Nature* **352**, p. 563

Daemmrich, A. A. (2014), “Investigating David Baltimore and Thereza Imanishi-Kari”, online article, SSRN, July 16, 2014. https://ssrn.com/abstract=2587959 or http://dx.doi.org/10.2139/ssrn.2587959

Doty, P. (1991), “Responsibility and Weaver et al.”, *Nature* **352**, pp. 183-184.

García, J.A.; Rodriguez-Sánchez, R. and Fdez-Valdivia, J. (2016), “Why the referees’ reports I receive as an editor are so much better than the reports I receive as an author?”, *Scientometrics* **106(3)**, pp. 967–986. https://doi.org/10.1007/s11192-015-1827-8

Gunsalus, C.K. (1999), "Book Review: The Baltimore Case: A trial of politics, science, and character, by Daniel J. Kevles", N. Engl. J. Med. **340**, p. 242, doi: 10.1056/NEJM199901213400320

Haffar, S.; Bazerbachi, F. and Murad, M. H. (2019), “Peer review bias: a critical review”, *Mayo Clinic Proceedings* **94(4)**, pp. 670–676.

Horbach, S. & Halffman, W. (2018), “The changing forms and expectations of peer review”, *Research Integrity and Peer Review* **3(8)**, https://doi.org/10.1186/s41073-018-0051-5

Horta, H.; Jung, J. (2024), “The crisis of peer review: Part of the evolution of science”, *Higher Education Quarterly* **78**, e12511, https://onlinelibrary.wiley.com/doi/full/10.1111/hequ.125114

Hosseini, M. and Resnik, D. B. (2024), “Guidance needed for using artificial intelligence to screen journal submissions for misconduct”, *Research Ethics* **21(1)**, https://doi.org/10.1177/17470161241254052

Huisman, J. and Smits, J. (2017), “Duration and quality of the peer review process: the author’s perspective”. *Scientometrics* **113(1)**, pp. 633–650. https://doi.org/10.1007/s11192-017-2310-5

Ibrahim, H.; Liu, F.; Zaki, Y., and Rahwan, T. (2025), "Citation manipulation through citation mills and pre-print servers", *Sci Rep* **15(5480)**, https://doi.org/10.1038/s41598-025-88709-7

Jackson, A. (2007), “Jumping Ship: Topology Board Resigns”, *Notices of the American Mathematical Society* **54(5)**, pp. 637-639.

Lang, S. (1993), “Questions of Scientific Responsibility: the Baltimore case”, *Ethics and Behavior* **3 (1)**, pp. 3-72.

Lang, S. (2000), “On a Yale Kevles Appointment", *Yale Daily News,* 3 February 2000, pp. 6-9, https://ydnhistorical.library.yale.edu/?a=d&d=YDN20000203-01.2.28&e=-------en-20--1--txt-txIN--------false#

Linse, A. R. (2017), “Interpreting and using student ratings data: Guidance for faculty serving as administrators and on evaluation committees”, *Studies in Educational Evaluation* **54**, pp. 94-106.

Loewe, B. (2022), "Measuring the Agreement of Mathematical Peer Reviewers", *Axiomathes* **32(3)**, pp. 1205-1219. https://doi.org/10.1007/s10516-022-09647-x

Moustafa, K. (2022), "No to paid peer review", *The Lancet* **400(10347),** Correspondence, p. 160, July 16, 2022

Nature (1991), “Even misconduct trials should be fair”, Editorial, *Nature* **350**, pp. 259–260. https://doi.org/10.1038/350259a0

Nature (2018), “Chinese checkers. China sets a strong example on how to address scientific fraud”, Editorial, *Nature* **558**, p. 162.

O’Toole, M. (1991), “Margot O’Toole’s Record of Events”, *Nature* **351**, pp. 180-183.

Parson, V. (2025), “Beyond penalties: a proactive approach to developing student understanding of academic integrity”, *Assessment & Evaluation in Higher Education* **50(4)**, pp. 592-606, https://doi.org/10.1080/02602938.2025.2580621

Rejecta Mathematica (2009), https://web.archive.org/web/20090803131854/http://math.rejecta.org/

de Russy, C. and Langbert, M. (2005), “The Corrosion of Ethics in Higher Education”, *Inside Higher Ed*, online article (July 4, 2005), https://www.insidehighered.com/views/2005/07/05/corrosion-ethics-higher-education

Scarcella, M. (2024), “Academic publishers face class action over ‘peer review’ pay, other restrictions”, *The Thomson Reuters*, online article, Sep. 13, 2024, https://www.reuters.com/legal/litigation/academic-publishers-face-class-action-over-peer-review-pay-other-restrictions-2024-09-13/

Sklamberg, H. (1990), “Education?”, *Yale Daily News,* 27 September 1990, pp. 5-6 (reprinted by S.Lang), https://ydnhistorical.library.yale.edu/?a=d&d=YDN19900927-01.2.56&e=-------en-20--1--txt-txIN--------false

Skórzewska-Amberg, M. (2022), “Publications in Scientific Journals – the New Face of Deception”, *Krytyka Prawa (The Critique of Law)* **14(4)**, pp. 20-28, https://doi.org/10.7206/kp.2080-1084.553

Stoimenov, A. (2022), "Conduct and Correctness in Mathematical Publishing", *Publishing Research Quarterly* **38**, pp. 586-598, https://doi.org/10.1007/s12109-022-09895-0

Stoimenov, A. (2023), "Reliability or Liability in the Contemporary Mathematics Publishing Process? An Ethical and Technological Case Study", *Cogent Social Sciences OASS* **9(2)**, 2244259, https://doi.org/10.1080/23311886.2023.2244259

Stoimenov, A. (2025), “Confidentiality, Peer Review, and the Submission Case Database”, *The Journal of Scholarly Publishing* **56(4)**, pp. 733-755. doi: 10.3138/jsp-2024-0093

Stoimenov, A. (2026), "Designing an Evaluation Process for Science Editors", available from https://stoimenov.net/stoimeno/homepage/papers.html

Weiss, P. (1989), “Conduct Unbecoming?”, *The New York Times*, October 29, 1989, https://www.nytimes.com/1989/10/29/magazine/conduct-unbecoming.html

Wolfram, D.; Wang, P.; Hembree, A. and Park, H. (2020), ″Open peer review: promoting transparency in open science", *Scientometrics,* **125(2)**, 1033–1051. https://doi.org/10.1007/s11192-020-03488-4

Zaharie, M.A. and Seeber, M. (2018), “Are non-monetary rewards effective in attracting peer reviewers? A natural experiment.”, *Scientometrics* **117(3)**, pp. 1587–1609, https://doi.org/10.1007/s11192-018-2912-6

Zhang, Y. (2024), “Who should be held accountable? An analysis of academic misconduct cases in Chinese higher education and the challenges of the transformative process”, *Region - Educational Research and Reviews* **6(5),** pp. 144--149. DOI: 10.32629/rerr.v6i5.2164