

8-16-1996

From Isolation to Collaboration: A WWW based Collaborative Review System

Ajaz R. Rana

New Jersey Institute of Technology, rana@cis.njit.edu

James Whitescarver

New Jersey Institute of Technology, jim@helios.njit.edu

Sudha Godala

New Jersey Institute of Technology, sunda@cis.njit.edu

Firas Aljallad

New Jersey Institute of Technology, firas@cis.njit.edu

Follow this and additional works at: <http://aisel.aisnet.org/amcis1996>

Recommended Citation

Rana, Ajaz R.; Whitescarver, James; Godala, Sudha; and Aljallad, Firas, "From Isolation to Collaboration: A WWW based Collaborative Review System" (1996). *AMCIS 1996 Proceedings*. 145.

<http://aisel.aisnet.org/amcis1996/145>

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISEL). It has been accepted for inclusion in AMCIS 1996 Proceedings by an authorized administrator of AIS Electronic Library (AISEL). For more information, please contact elibrary@aisnet.org.

From Isolation to Collaboration: A WWW based Collaborative Review System

[Ajaz R. Rana](mailto:rana@cis.njit.edu) (rana@cis.njit.edu), [James Whitescarver](mailto:jim@helios.njit.edu) (jim@helios.njit.edu), [Sudha Godala](mailto:sudha@cis.njit.edu) (sudha@cis.njit.edu), [Firas Aljallad](mailto:firas@cis.njit.edu) (firas@cis.njit.edu)
New Jersey Institute of Technology

Abstract: Decisions to accept or reject manuscripts when referees have polarized opinions about the worth of a manuscript entreat the scientific community to devise means that discourage the use of subjective criteria and increase the level of inter-referee agreement. This paper describes a decision model and a system that not only provides a means to improve the inter-referee agreement, but also includes mechanisms for curtailing often noted group process losses. It is argued that the World Wide Web based Distributed Group Support System allows opportunities for implementing and experimenting with various communication and decision structures that have significant ramifications for the peer review process.

1. Introduction: The controlled dissemination of scientific knowledge has implications for both scientific knowledge and scientists' careers. Oft-noted pitfalls in the referee system have led to various proposals for improving the system (Peters & Ceci, 1982; Kupfersmid, 1988; Chubin & Hackett, 1990; Cicchetti, 1991; Scott, 1974; Glenn, 1976; Harnad, 1994; and others). Despite the adoption of various quality control mechanisms, such as double blind refereeing, the decision making process leaves much to be desired.

It is documented that variations in decision making procedures exist among journals and fields of inquiry (Hargens, 1988). Besides the variations in decision procedures, it is postulated that prevailing norms regarding the risk of rejecting a worthy manuscript and of accepting a sub-standard manuscript, differ from one field of research to the other (Zuckerman & Merton, 1971). The relationship of normative and structural differences among journals and fields of inquiry is debated to have an impact on the following two, rather disturbing factors: (i) low levels of inter-referee agreement, and (ii) high rejection rates. Several researchers have reported low inter-rater agreement for various social and behavioral science journals (Munley et al, 1988; Whitehurst, 1984; Gottfredson, 1978; Smigel & Ross, 1970; Scott, 1974). The inter-referee agreement (or inter-reliability) coefficients, as determined by intra-class correlation for some of the social science journals, are as low as 0.20. The high rejection rates in social sciences (as high as 80% for some journals) combined with low reliability figures portray a very flimsy picture of the traditional review process. What causes these low levels of agreement? Cole, et al (1981), suggest that a large proportion of reviewer disagreement may be a result of real and legitimate differences of opinion among experts about what good science is or should be. Others suggest that it is biases (such as gender, institutional affiliation, replication, self-citations, statistical significance, etc.), use of subjective criteria, and/or errors of omission that underlie such low reliabilities (Cicchetti, 1991). Whatever the case may be, it is clear that the referee system needs improvement.

2. A DGSS based Review Model: Rana (1995) proposed that "after a usual review of the manuscript and before the editorial decision, the reviewers should have the opportunity to share their reviews and resolve the areas of wide disagreement, especially on the points where their judgments are polarized." (pp. 37). This review model differs from the traditional review process where referees complete their reviews in isolation from each other and hence provides no opportunity for resolution of disagreements and assimilation of judgments before the editorial decision. The theoretical foundations of this decision making model are rooted in research with small decision making groups. Various mechanisms demonstrated to have a positive effect on group process and outcome, such as NGT, Delphi, and SJT (Gustafson et al., 1973; Van de Ven & Delbecq, 1974; Dalkey & Helmer 1963; Linstone & Turoff, 1975; Roharbaugh, 1979; 1981), share a common approach. That is, the quality of group work (on a variety of tasks) can be improved if group members engage in a discussion of the task at hand, having first provided their individual responses.

Rana and colleagues (Rana, Turoff, & Hiltz, 1996) further argue that the proposed review model can be implemented via the use of Distributed Group Support Systems (DGSS). They note the following reasons for the suitability of DGSS as opposed to decision room based Groups Support Systems: (i) a DGSS would not disallow any of the reported or commonly used mechanisms for improving the peer review process; and

(ii) the referees would have complete time and space freedom to engage in the review process. Rana (1995) implemented the above review model in a DGSS, called EIES2. Research with the use of this system showed that the controlled sharing of reviewer responses, was instrumental in enhancing the quality of the reviews, especially when reviewers had higher levels of pre-discussion disagreement (Hiltz, et al., 1996). The software feature intended to enhance the agreement among reviewers also had a positive effect on post-discussion agreement. However, the effect of the consensus enhancing tool was not statistically significant. The major conclusion of this study was that the strength of the DGSS based review process lies in its ability to allow for disagreement among reviewers before the discussion phase and later provide a mechanism for reducing the level of disagreement. The version of EIES2 used in this study provided a menu driven VT100 terminal user-interface, as opposed to a generally preferred graphical user interface. Based on the lessons learned from this study, the authors have built a WWW-based subset of EIES2 tailored for peer review purposes. At present, the system is in quality assurance stage and is accessible through WWW at <http://eies.njit.edu:2134>. The following section provides a brief description of this system.

3. System Design and Functionality: The Review System uses a Smalltalk-80 distributed object base to implement a user interface and communication structure tailored to conducting the review of manuscripts. For every submission, an editorial office can set up a shared discussion space, called a review group (or conference), for selected reviewers. The editorial office can set up an agenda for reviewers. The agenda can consist of various decision support tools. However, as suggested by the proposed review model, the inclusion of the following four steps comprise the minimum desirable actions that should be undertaken by the reviewers: (i) critical evaluation of the manuscript, (ii) rating the quality of the manuscript on provided scales, (iii) providing justifications for the ratings, and (iv) engaging in discussion to resolve areas of disagreement with other review members.

The first step, critical evaluation of manuscript, may be done off-line using a hard copy of the manuscript. However, the manuscript can be provided in electronic form (eg., as an HTML, ASCII, PostScript, or MS Word document, etc.) as well. The review system provides software support for rating the quality of the manuscript. The choice of the style and number of rating scales is completely in the hands of person setting up the agenda (in this case the editorial office). Once a reviewer has rated the paper on the requisite scales he/she can view the positions of others in the group. A summary of group members' ratings allows the reviewers to clearly see deviant positions. The provision of justification for ratings is done in the form of text items. Similar to ratings, no group members can see others' responses unless they have done their own responses. Reviewers may be required to provide justifications for their ratings on each individual scale separately, or as one composite item providing reasons for ratings on all the scales.

Once the users have completed the above steps, they engage in discussion to reach agreement on the value of the manuscript. At any time during the review process, reviewers can update their responses to the scales, indicating a change in position. Every argument or discussion point, called a comment, can be labeled, indexed, and linked with other comments. Such an arrangement makes it easy to capture and retrieve a chain of comments relating to a certain topic. Indexes are also maintained for every member of the group, representing the unread comments by a particular member.

4. Publication Decision Process and Research Opportunities: We believe, the state of the referee system has much to gain by adopting the proposed review system. The use of computer and communication technology to introduce the controlled sharing of individual reviews and the opportunity for resolution of disagreements offer a real possibility for achieving high reliability (high levels of inter-referee agreement) and validity (high quality) in reviews.

Persuasive Argumentation: It is hypothesized that the agreement among reviewers would result from persuasive argumentation, as opposed to normative influences or generally noted pressures toward conformity. If reviewers know that their opinions will be read and/or possibly criticized by other reviewers, they will be much more careful in their initial reviews. Additionally, since the reviewers, having committed to their positions about the manuscript, would enter the discussion phase with high valences (Hoffman, 1979) for their choices, they would be compelled to defend their positions. Such argumentation would no doubt result in bringing omissions to the foreground and fend off reviewer biases. In view of the legitimate and real differences in opinions about what good science is and should be, the objective of the review

system is not to achieve optimum/perfect agreement on the worth of the manuscript. Rather, the objective is to correct errors of omission and subjective judgments; and improve the level of agreement through argumentation, especially on those aspects of the manuscripts where referee opinions are on the opposite extremes.

Reviewer Identities: Every reviewer can be assigned a pen name (other than the user's real name). The reviewers may choose to reveal their real name, use the assigned pen name, or contribute under complete anonymity. In GSS research, the impacts of anonymity have been studied on various group process and outcome related variables (quality of outcome, post-discussion agreement, deindividuation, disinhibition, satisfaction, etc.). The current understanding seems to be that "the importance of anonymity depends on the nature of the group and the task" (Dennis & Gallupe, 1993; pp 74). Even though by definition the members of a research community acting as referees form a group of peers, a group of reviewers in reality consists of members with well recognized status differences. In an asynchronous DGSS based review process, we expect that the benefits of anonymity (critical evaluation, argumentation, reduced pressures toward conformity, etc.) would outweigh its costs (e.g., less likelihood of reaching consensus).

Having More Reviewers: Statistically, as long as the reviewer errors are independent (unbiased), the larger the group size the greater the reliability of reviews (Kraemer, 1991). However, an increase in reliability of reviews does not guarantee increase in validity, especially when experts are known to have biased opinions. Yet, with a larger group of reviewers, the chances of a major flaw in a manuscript going unnoticed are reduced, and hence higher validity is achievable. Both experimental and field research with GSS also shows that the larger groups gain more from GSS use than do the smaller groups (Dennis & Gallupe, 1993). As a result, we expect that a larger than usual group size (3 person) would be better suited to an asynchronous DGSS review process. Following is an outline of some of the alternative communication and publication decision structures that can be implemented via a DGSS.

Simultaneous Rating of Sets of Papers: It has been suggested that simultaneous rating of sets of papers can improve inter-rater reliabilities (McReynolds, 1971; Perlman, 1982). The proposed collaborative review process allows the possibility of experimenting with such simultaneous comparisons.

Author's Response before Editorial Decision: The DGSS based review model provides efficient means of implementing the proposal that authors should be able to respond to the reviewers' critique before the disposition decision (Glenn, 1976).

The Adversary Model: Bornstein (1990) suggests that each manuscript should be considered publishable unless shown by the reviewers (the prosecuting attorneys) to be seriously flawed. In this decision model the author plays the role of defense attorney.

Successive Proportional Additive Numeration (SPAN): The proposed review model allows the possibility of implementing SPAN voting process (MacKinnon & MacKinnon, 1969; Frost & Taylor, 1985). This scheme works by giving each reviewer a fixed number of votes which he/she may allocate to another reviewer or the manuscript to be reviewed. This process is repeated until all votes are distributed to a decision choice about the manuscript.

Review Group Membership: The membership of a review group (or conference) can be limited to the reviewers and the editors. It is also possible to make the conference public to the rest of the user community, but restrict their privileges as mere observers.

5. Conclusion: The vital issues for the proposed mode of reviewing are (i) will the WWW based review system actually be adopted by the actors in the referee system (editors, reviewers, & authors)? and (ii) if adopted, will the technological features be appropriated by the reviewers in effective manner (DeSanctis & Poole, 1994)? The most prudent answers to these questions rest upon, yet to be conducted, field trials of the proposed review system.

In the absence of the representative data on the DGSS supported review process, the theory of critical mass (Markus, 1990) suggests that universal access to the review system would be crucial for the adoption and success of the alternative review process. Universal access or the ability of any member of scientific community to reach all other members through the WWW is important because without universal access, the referee system faces the risk of isolating some preferred referees and shrinking the pool of reviewers. The recent proliferation of nodes on the Internet and the ownership of personal computers leaves no room for skepticism in this regard. The theory of critical mass further suggests that a sizable number of influential initial users (critical mass) would be pivotal for achieving universal access. Whether critical mass for the proposed system constitutes a community of influential scientists (essentially the editors and the reviewers of prestigious journals) remains to be seen. Regarding the appropriation of the system features, the results of the controlled experimental study conducted at NJIT (Rana, 1995) provide support for the predictions of the theory of adaptive structuration (DeSanctis & Poole, 1994). That is, the social construction of the technology in reviewer groups would depend greatly upon the reviewers' attitudes towards the technology (DeSanctis & Poole, 1994).

We close this discussion by highlighting the need for field research. The IS community has already taken a leadership role in adopting innovative technologies to improve the dissemination of scientific knowledge (Watson, 1994). An initiative in exploring the potential of the proposed review system will enrich the electronic field of Intellectual Infrastructure and further strengthen the position of IS community.

Acknowledgments: This research was partially supported by grants from the National Science Foundation on Coordination Theory and Collaboration Technology (NSF-IRI-9015236 and NSF-IRI-9408805) to Roxanne Hiltz and Murray Turoff. The opinions expressed do not necessarily represent those of the National Science Foundation. Among the many people who have contributed to the program of research are: Marsha Roberts; Srinivas Reddy; Robert Czech; Jerry Fjermestad; Ronald Rice; Scott Poole; and Richard Hoffman.

References: Available from authors.

References:

Armstrong, J.S. (1982a) Barriers to Scientific Contributions: The authors formula. *The Behavioral and Brain Sciences* 5:197-99.

Armstrong, J.S. (1982b) The Ombudsman: Is peer review as fair as it appears?. *Interfaces* 12:62-74.

Bakanic, V., McPhail, C. & Simon, R.J. (1987) The Manuscript Review and Decision-Making Process. *American Sociological Review* 52:631-42.

Bornstein, R. F. (1990) Manuscript Review in Psychology: an Alternative Model. *American Psychologist* 45:672-73.

Chubin, D.E. & Hackett, E.J. (1990) *Peerless Science : Peer Review and V.S. Science Policy*. State University of New York Press.

Cicchetti, D.V. (1991) The Reliability of Peer Review for Manuscript and Grant Submissions : A Cross-Disciplinary Investigation. *Behavioral and Brain Sciences* 14:119-86.

Cole, S., Cole, J.R. & Simon, G.A. (1981) Chances and Consensus in Peer Review. *Science* 214:881-86.

Crandall, R. (1991) What should be done to improve reviewing? *Behavioral and Brain Sciences* 14:143.

Cummings, L.L. & Frost, P.J. (1985a) *Publishing in the Organizational Sciences*. Richard D. Irwin, Inc.

Dalkey, N. & Helmer, O. (1963) An Experimental Application of the Delphi Method to the Use of Experts. *Management Science* 9:458-67.

DeScantis, G.L., & Poole, M.S. (1994) Capturing the Complexity in Advanced Technology Use: Adaptive Structuration Theory. *Organization Science*. 5:2:121-147.

Dennis, A. R. & Gallupe, R. B. (1993) A History of Group Support Systems Empirical Research: Lessons learned and Future Directions. In: *Group Support Systems: New Perspectives*, Ed. L. M. Jessup & J. S. Valacich. Maxwell Macmillan International.

Frost, P. J. & Taylor, R. N. (1985) Partisan Perspective: A Multiple Level Interpretation of the Manuscript Review Process in Social Science Journals. In *Publishing in Organizational Sciences*, Ed. L. L. Cummings & P. J. Frost. Richard D. Irwin Inc.

Garvey, W.D., Gottfredson, S.D. & Simmons, J.G. (1984) A Comparison of Two Major Scientific Information Exchange Processes in Psychology: 1962 and 1976. *American Psychologist* 39:11-21.

Glenn, N.D. (1976) The Journal Article Review Process : Some Proposals for change. *The American Sociologist* 11:179-85.

Gottfredson, S.D. (1978) Evaluating Psychological Research Reports: Dimensions, Reliability, and Correlates of Quality Judgments. *American Psychologist* 39:22-28.

Gustafson, D.H., Shukla, R.K., Delbecq, A. & Walster, G.W. (1973) A Comparative Study of Differences in Subjective Likelihood Estimates Made by Individuals, Interacting Groups, Delphi Groups, and Nominal Groups. *Organizational Behavior and Human Performance* 9:280-91.

Hargens, L.L. (1988) Scholarly Consensus and Journal Rejection Rates. *American Sociological Review* 53:139-51.

Harnad, S. (1979) Creative Disagreement. *The Sciences* 19:18-20.

Harnad, S. (1994) Implementing Peer Review on the Net: Scientific Quality Control in Scholarly Electronic Journals. *Proceedings of International Conference on Refereed Electronic Journals: Towards a Consortium for Networked Publications*. University of Manitoba, Winnipeg 1-2. October (1993) (in press).

Hiltz, S. R., Dufner, D, Fjermestad, J., Ocker, R., Rana, A.,R. & Turoff, M. (1996) Distributed Group Support Systems: Theory Development and Experimentation. Book chapter: Olsen, B. M., Smith, J. B. & Malone, T., eds., *Coordination Theory and Collaboration Technology*, Hillsdale NJ: Lawrence Erlbaum Association, 1996.

Hoffman, L.R. (1979) *The Group Problem Solving Process: Studies of Valence Model*. Praeger, New york.

Kupfersmid, J. (1988) Improving What is Published: A Model in search of an Editor. *American Psychologist* 43:635-42.

Kraemer, H. C. (1991) Do we really want more "reliable" reviews? *Behavior and Brain Sciences* 14:152-54.

Linstone, H. & Turoff, M. (1975) *The Delphi Method : Techniques and Applications*. Addison Wesley Publishing Company.

- MacKinnon, W. J. & MacKinnon, M. M. (1969) Computers in Behavioral Science: The Decisional Design and Cyclic Computation of SPAN. *Behavioral Science* 14:244-47.
- McReynolds, P. (1971) Reliability of Ratings of Research Papers. *American Psychologist* 26:400-01.
- Mahoney, M.J. (1977) Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System. *Cognitive Therapy and Research* 1:161-75.
- Mahoney, M.J. (1978) Publish and Perish. *Human Behavior* 7:38-41.
- Mahoney, M.J. (1985) Open Exchange and Epistemic Progress. *American Psychologist* 40:29-39.
- Markus, M.L. (1990) Toward a "Critical Mass" Theory of Interactive Media. In: *Organizations and Communication Technology*, Ed: J. Fulk and C. Steinfield, PP.194-218. Beverly Hills, CA: Sage.
- Munley, P. H., Sharkin, B., & Gelso, C. J. (1988) Reviewer Ratings and Agreement on Manuscripts Reviewed for the *Journal of Counseling Psychology*. *Journal of Counseling Psychology* 35:2:198-202.
- Perlman, D. (1982) Reviewer "bias": Do Peters and Ceci protest too much? *The Behavioral and Brain Sciences* 5:231-32.
- Peters, D.P. & Ceci, S.J. (1982) Peer - review Practices of Psychological Journals : The Fate of Published Articles, Submitted Again. *The Behavioral and Brain Sciences* 5: 187-255.
- Rana, A. R. (1995) Peer Review Process and Group Support Systems: Theory Development and Experimental Validation. Ph.D. Dissertation, Rutgers University.
- Rana, A. R., Turoff, M., & Hiltz, S. R. (1996) Peer Review Process: A Group Support Systems Approach. Submitted for publication (under review).
- Rathie, S. E. (1996) Electronic Journals and Peer Review: Perils and Promises. URL <http://www.cs.ubc.ca/spider/rathie/elecpub/paper3.html>
- Rohrbaugh, J. (1979) Improving the Quality of Group Judgment: Social Judgment Analysis and the Delphi Technique. *Organizational Behavior and Human Performance* 24:73-92.
- Rohrbaugh, J. (1981) Improving the Quality of Group Judgment: Social Judgment Analysis and the Nominal Group Technique. *Organizational Behavior and Human Performance* 28:272-88.
- Rourke, B.P. (1991) Toward Openness and Fairness in the Review Process. *Behavioral and Brain Sciences* 14 : 161.
- Scott, W.A. (1974) Interrefree Agreement on some Characteristics of Manuscripts submitted to the *Journal of Personality and Social Psychology*. *American Psychologist* 29:698-702.
- Smigel, E.O. & Ross, H.L. (1970) Factors in the Editorial Decision. *The American Sociologist* 5:19-21.
- Spencer, N.J., Hartnett, J. & Mahoney, J. (1986) Problems with the Reviews in the Standard Editorial Practice. *Journal of Social Behavior and Personality* 1:21-36.
- Van de Ven, A. & Delbecq, A.L. (1974) The Effectiveness of Nominal, Delphi, and Interacting Group Decision Making Processes. *Academy of Management Journal* 17:605-21.

Watson, R. (1994) Creating and Sustaining a Global Community of Scholars. *MIS Quarterly*, September 225-231.

Whitehurst, G.J. (1984) Interrater Agreement for Journal Manuscript Reviews. *American Psychologist* 39:22-28.

Zuckerman, H. & Merton, R.K. (1971) Patterns of Evaluation in Science: Institutionalization, Structure, and Functions of the Referee System. *Minerva* 9:66-100.