Title

A normative model of peer review: qualitative assessment of manuscript reviewers' attitudes towards peer review

Authors

Tercier, John
Callaham, Michael L.

Publication Date

2007-11-01

Title: A normative model of peer review: qualitative assessment of manuscript reviewers' attitudes towards peer review

Manuscript Type: Original Research

Section/Category: Education

Keywords: peer review, publication, education

Authors: John Tercier, MD, PhD
Department of Anthropology, History and Social Medicine, University of California San Francisco, 3333 California St., San Francisco, CA, 94143-0850

Michael L. Callaham, MD
Division of Emergency Medicine, University of California San Francisco, Box 0208 , San Francisco CA 94143-0208
Michael.Callaham@ucsf.edu

Competing interests: None. No external funding or support was received for this research.

Presented at the Fifth International Congress on Peer Review and Biomedical Education, Chicago IL 9/18/05 (see http://www.ama-assn.org/public/peer/abstracts.html#predict2, Saturday poster section, "Peer Review Process", go to "Qualitative Profile of Journal Peer Reviewers and Predictors of Peer Reviewer Quality" and click on "abstract".)

Manuscript completed 11/07

ABSTRACT

Objective: Peer review is considered crucial to the selection and publication of quality research, yet little is known of the values, beliefs and attitudes of peer reviewers towards the process of peer review. This study elicits reviewer beliefs about the process in order to produce a normative model of peer review.

Methods and Findings: The 72 subjects were experienced reviewers at *Annals of Emergency Medicine* and had completed at least 5 reviews in the past 2 years. Subjects participated in 40 minute structured telephone interviews focusing on reviewer attitudes, beliefs and values towards the process of peer review. Subject responses were coded and categorized using grounded theory to produce a qualitative profile of reviewers' attitudes towards peer review and generate a normative model of the peer review process. This model was found to closely adhere to conventionally held beliefs

about the process of peer review. However, within it were revealed a number of areas where reviewers, aware of tensions within the process, questioned those conventional beliefs, expressing concern about methods, operations and outcomes. As researchers producing research and receiving reviews and as reviewers judging others' research and producing reviews, the reviewer's status as "peer" was seen as both essential to the operation of the system and problematic. In their perception of the role of the peer reviewer, though respondents identified evaluation of the manuscript (selecting submissions for publication by filtering out incorrect or inadequate work) as the primary goal of the formal process, instruction of the researcher (improving the accuracy, clarity and utility of published research) was considered by the majority of reviewers as the more important practice. Likewise, though there was recognition that the review process aims to prevent poor research from being published, there was more concern over the danger that it results in good research being "strangled in its cradle". Though respondents believed that the quality of the review is determined primarily by the skills of the individual reviewer, they maintain that the validity of the process is determined by the corporate nature of the review panel acting as a system of checks and balances. Though perceiving this system of checks and balances as requiring a degree of separation of authorial, review and editorial functions, reviewers, at the same time, express the desire for a more open system of feedback leading towards a more consensual research outcome. Two issues of concern arose repeatedly in the interviews: frustration at a perceived lack of feedback from editors to reviewers and repeated condemnations of "mean-spirited" feedback from reviewers to authors. Defects in feedback were cited by respondents as a major barrier to optimizing research quality and editorial judgment.

Conclusions: The tensions found in the peer review process, sometimes seen as barriers to its effective operation, are less defects in the process than definitive of the concept of peer review itself and thus necessary to its operation. While at a practical level peer review operates as a triage exercise, it is, at the same time, on a social level, a mode of disciplinary dialogue between peers: important not only to the maintenance of an effective knowledge base and thus disciplinary validity, but also, through its effect on researchers and reviewers, important in the construction of disciplinary identity. Peer review's practical and social operations are not antithetical to each other but rather are inherent in the hybrid concept of the peer reviewer, where one's status as a peer makes possible one's activity as reviewer.

Title: A normative model of peer review: qualitative assessment of manuscript reviewers' attitudes towards peer review

BACKGROUND

Given the importance of peer review within contemporary science and medicine, it is surprising, as a number of authors have pointed out; that it remains so poorly understood (Callaham 2002, 313; Ingelfinger 1974, 686; Kassirer 1994, 96; Rennie 1988, 181). The research that has been done on the sensitivity, specificity and reliability of the manuscript peer review process has up to now yielded disappointing results (Baxt 1998, 310; Cole 1981, 881; Goodman 1994, 11; Pierie 1996, 1480; Rothwell 2000, 1964). Attempts have been made by editors to improve the efficacy of the process by improving the training and selection of peer reviewers, but studies measuring the effect of such interventions on review quality have, again, yielded disappointing results (Bazarian 1999, 148; Callaham 1997, 192; Callaham 1998, 318; Callaham 2002, 323; Garfunkel 1990, 1369; Langkamp 1992, 528; Linzer 1988, 2537; Norman 1998, 158 ; Schroter 2004, 637; Stossel 1985, 658; Strayhorn 1993, 947). The process of peer review appears to be such an integral, logical and thus "natural" part of the scientific process that its structure, function, method and value are assumed to be self-evident. The mechanics of the process are given in journals' instructions to reviewers and there are no lack of editorial pieces stating what peer review *ought* to be (Arrington 1995, 249; Jefferson 2002, 2784; Wager 2001, 257; Rennie 1998, 300; Smith 2006, 178). However, it is not known to what extent the concepts, goals, methods and assessment criteria advocated by editors are accepted by reviewers, nor what reviewers' attitudes are towards the "standard of care" expected of them. This study constructs a normative model of manuscript peer review based on a qualitative profile of reviewers' beliefs about the review process. That is, it describes a model of peer review as it is conceived of by peer reviewers.

METHODS

Setting: Manuscript reviewers at *Annals of Emergency Medicine*.

Selection of participants: *Annals of Emergency Medicine* is the leading journal in the specialty of emergency medicine and ranks in the top 11% among 5,876 science and medical journals listed by the ISI in frequency of citations (Journal 2004). *Annals* has a pool of 487 active regular reviewers (one time guest reviewers were omitted). The

sample used in this study was limited to those reviewers who had completed at least 5 reviews in the past 2 years. Reviewers who met those criteria had their identity concealed and were then stratified to represent a broad cross section of relevant characteristics, including quality of reviews (using a standardized 5 point score), years of experience peer reviewing for any journal, academic rank, experience on editorial boards, additional training in epidemiology or statistics, and experience on grant review panels. The resulting sample of 107 reviewers represented maximal heterogeneity of qualifications, and were contacted by e-mail to request their participation. Non-respondents were sent two follow-up e-mails. Failure to respond to these attempts resulted in elimination from the study. Of the 107 subjects contacted, 72 responded and were interviewed.

Data collection and processing: An interview template for a prospective qualitative semi-structured telephone interview as per Fowler (1995) and Silverman (1997, 2001) was designed by consensus of two investigators (JT, MC). It was piloted on four reviewers taken at random from the respondent pool to assess for understandability and practicality, and underwent revisions based on this process. All interviews were carried out, recorded and transcribed by a single investigator (JT) to ensure consistency. A standard introductory paragraph outlining the format of the interview was read to each participant. Participants were informed of the data collection process, assured that responses would not be attributed to individuals, and permission was requested to record the interview. There were no refusals from the group who agreed via e-mail to the interview.

Each participant was asked the same series of closed and open-ended questions according to the interview template (Appendix One). Interviewees were encouraged to elaborate on answers to open-ended questions. Few required such encouragement; most respondents were eager to enter into a dialogue about the process. Answers were explored in greater or lesser degrees for clarification and understanding at the discretion of the interviewer using predetermined qualifier questions. Responses were recorded and written down on a standardized form as direct quotes in verbatim fashion. The response forms and recordings were retained as a permanent record of the conversation. Responses were then directly transcribed immediately after each interview from the standardized form into a composite transcript for analysis.

The transcripts were analyzed using qualitative grounded theory as described by

Glaser (1998), Strauss (1990) and Clarke. (2005) They were reviewed by the investigator, who categorized ideas and derived a code in typical grounded theory manner. Coding resulted in sets of general categories and specific qualifiers. Each novel idea was assigned a new strategy code, and in the case of subtle variation, a new qualifier code. When a new code was added to the list, the entire transcript was reviewed to reclassify ideas as needed. Each of the final transcripts were then reviewed using the final coded categories and recoded. Results were then reported using the normative categories developed from reviewers' responses. These are discussed below along with the inclusion of respondent comments and insights.

LIMITATIONS

Sample size and make-up: In quantitative social science research, a sample size in the region of 60 to 80 is considered the minimum required to pick up "large" effects. However, in qualitative research, where the goal is not prediction based on statistical probability but critical analysis through logical inference such sample sizes are considered large.The participants were heterogeneous in terms of experience, practice environment and quality rating scores, and an attempt was made to compensate for this through the stratification process utilized in participant selection. However, the subjects were reviewers successful enough to be retained by the journal for at least the two prior years and thus not a representative sample of all those who might initially be involved in review. Though we had a high response rate (67%), we speculate that reviewers who did not bother to respond might also be less experienced and less capable reviewers, so our study population may under-represent the opinions of that group. The respondents in this study all came from a single specialty peer review journal (as in previous studies), but had appointments at medical schools across the U.S. and a broad variety of backgrounds and training. This particular reviewer population has been well studied in the past, and has performed similarly to journal reviewers from other specialties. Because this reviewer population has been thoroughly compared to those of other journals, we believe that our results are likely to be generalizable to peer reviewers in other specialties and journals (Baxt 1998, 310; Callaham 1997, 192; Callaham 1998, 318; Callaham 2002, 323; Callaham 2002, 2781; Nylenna 1994, 149; Schroter 2004, 673).

Interview situation: Interviews are a set of social interactions, whereby the response is dependent upon: the context of the interview, the interview's structure, and the interviewer's questions, affect and responses. Implicit in qualitative social science research that relies upon the subjects' reporting and interpretation of their own actions is that respondents do not necessarily report their behavior reliably. Even in the absence of any direct intent to mislead: respondents may forego reporting what they think is irrelevant, may "second-guess" the intent of the interview and may respond in ways structured by the interview situation. Nevertheless, respondent statements can be clarified and explicated by careful discussion between the interviewer and respondent and by cross-checking with complementary questioning in different sections of the interview. Another limitation arises from the possibility that recorded interviews may lead to less candid responses and social expectation biases. However, the impression of the interviewer was that after the first few minutes of preliminary demographic data gathering, the presence of the recording device no longer registered with the respondents.

Qualitative interpretation: The interpretation of empirical data in qualitative research is achieved through logical (rather than statistical) inference making use of relevant theoretical principles (Lambert 2002, 210). The nature of qualitative research is such that it requires large amounts of interpretation on the part of the interviewer and the inevitable bias that intervenes must be taken into account by the reader.

ETHICAL REVIEW

Respondents were assured that all comments would remain anonymous and that specific comments would not be attributed to individuals in publication of the data. Respondents were told that participation in the interview would be considered implied consent and all demonstrated understanding of and agreement with this. This project received institutional ethics review board approval.

Competing interests: None. No external funding or support was received for this research.

RESULTS

I. Respondent Characteristics

The interview was designed to be 40 minutes long. The actual mean length of interview was 43 minutes with a range of 20 to 90 minutes (median, 40; SD, 14.27). The mean number of years that respondents had been reviewing was 9.33 (median, 7; SD, 4.77), with a range of 2 to 20 years. The average number of reviews done per year for *Annals* was 6 (median, 6; SD, 3.2). However, 86.8% (59/68) of reviewers reviewed for other journals, so that the average total number of reviews total done per year was 13 (median, 12; SD, 6.69) with a minimum of 2 and maximum of 45. An average of 13 reviews per year coincides with reviewers' assessment of 10 to 12 reviews as an appropriate number of reviews to do per year. The journals, other than *Annals*, for which reviewers most frequently reviewed were *Academic Emergency Medicine* (55.9%, 38/68), *JAMA* (17.6%, 12/68) and *Journal of Emergency Medicine* (11.8%, 8/68). 56.9 % (37/65) of respondents were on an editorial board. 45.9% (17/37) of those were on the editorial board of *Annals*. The mean number of years respondents had been editors was 3.85 (median, 3; SD, 2.4).

II. The Mechanics of Reviewing

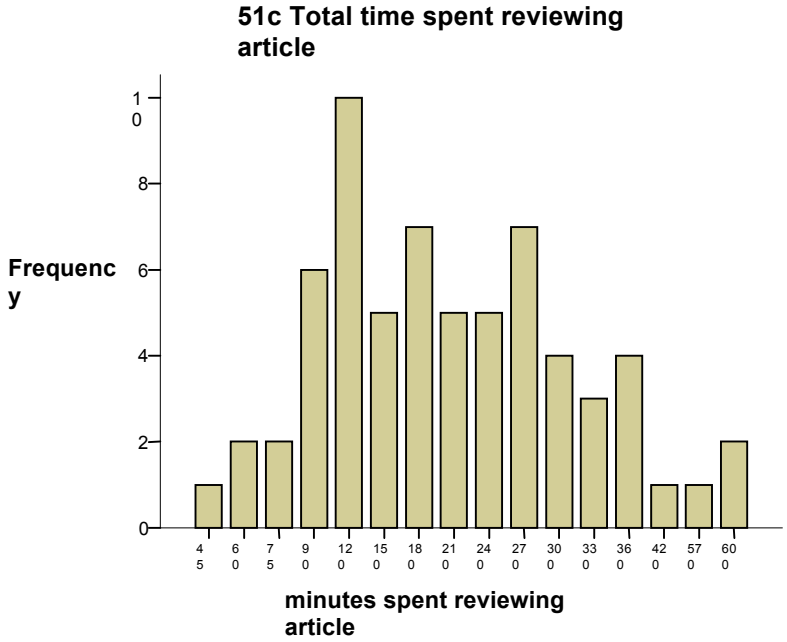
Respondents accept manuscripts based on their ability to clear space in their schedules. 80% stated that they do the review within 7 days of receiving the manuscript. 41% of respondents do their reviews at home, 22% do them at work, and 37% do them whenever they have time while at work, home or traveling. The location where reviews were done was determined by two things: a solid block of undisturbed time and access to the internet.

80% of respondents read the paper at least 2 or 3 times and 92% of those did it in 2 or 3 separate sittings. The majority of these reviewers stated that they liked to let a certain amount of time pass between readings—sometimes as little as an hour or two, other times one or two days.. A typical pattern consisted of:

> *'I read it through once, writing on it. Set it aside for a couple of days, then read it through again more slowly, doing a line by line reading, making detailed comments. Set it aside again for a couple of days, then come back and look at general things and write my comments to the editor and author' (37).*

The first read by the reviewer is described as a quick pass over the paper to get a general idea of what the research is about and whether they feel comfortable reviewing

it. The second read is to pick up and comment on “big picture” items, looking for major errors that can’t be fixed and less fatal errors that require fixing. The third pass is to do a detailed line by line critique and, if necessary, revise the reviewer's prior general opinion. There was, of course, some variation in this pattern, with many respondents combining aspects of the second with the first read, or the third with the second. Nevertheless, the consistency of the above process as described by respondents was remarkable. The most significant variation in review mechanics was in the amount of time spent doing the review. The mean time spent on doing a review was 3 hours 35 minutes, with a range of 1 to 10 hours (median 3, SD 2.05).



The structure of the reviewing process paralleled the structure of the manuscript, both in terms of the categories of analysis used and the order in which reviewers dealt with them. Respondents’ descriptions of the mechanics of review were so extremely uniform that it quickly became apparent that any attempt to quantify differences in procedure would be meaningless. The following composite description of the procedure of review is also a typical description:

ABSTRACT
- What is the manuscript about?
- Is it in my area of expertise?

INTRODUCTION
- Is it concise?
- Does it tell you what the study is about?

- Does it make you interested in the study?

HYPOTHESIS
- What is the study question?
- What are the stated goals of the study?
- Does it answer the “So what?” question?

METHODOLOGY / STUDY DESIGN
- Is the study design valid and the methodology appropriate?

RESULTS
- Are the results self-explanatory and clearly presented?

DISCUSSION
- Does the discussion deal with the data as it is presented?
- Does the discussion place the study within the relevant literature?

CONCLUSION
- Does the conclusion answer the question that was asked in the hypothesis?
- Have the stated goals been accomplished?
- Is the conclusion supported by the results?
- Does the conclusion avoid overreaching extrapolation?

GRAPHS / TABLES
- Are graphs and tables clear and comprehensible?

CITATIONS
- Are all the important articles about the subject mentioned?

All respondents stated that they preferred to print out and work off of a hard copy, with most making marginal notes on that hard copy. They then go on to formulate a list of bullet points from the marginal comments. These are transposed to the computer, at the same time turning the bullet points into a narrative. The resulting notes are used to compose the letter to the editor and comments to the author. The letter to the editor is usually written first and typically consists of a short summary of the paper to show the editor that the reviewer understands what the paper is about. This is followed by a statement about the importance of the issue being discussed, its clinical and scientific interest, then a list of major strengths and flaws. 48.5% (33/68) of respondents then give a recommendation to publish or not to publish and reasons for that recommendation. The letter to the author is then written as a revision of the editor’s letter in which the language has been reconfigured less judgmentally and more constructively. The majority of respondents say that they like to start off the comments to the author with positive points, then go on to outline major concerns, then a detailed line by line critique.

Despite the uniformity of process as described by respondents, most (85.3%, [58/68]) denied routine use of a formal template or checklist, though many pointed out that the journal's standard submission format acts as a kind of template when it comes to writing up the review. Clinical reference books were referred to only very rarely (an average of 4.39% of reviews [median, 1; SD, 6.94; range, 0 to 30%]). References on

study design or statistics were used more frequently (an average of 13.71% of reviews [median, 5; SD, 20.75; range, 0 to 90%]). Most reviewers commented that though they leave the statistics up to the statistical reviewers, they do make a "rough" attempt to assess the validity of the statistics themselves. A number of respondents commented that the presence of the statistical reviews gave them more confidence in their content review.

70% of reviewers "eyeball" citations 100% of the time. The reasons given for doing this were: in order to see that major studies in the area are present, that the citations are up to date, and that text and citations correspond. Reviewers pulled abstracts or manuscripts, on average, 40% of the time (median, 30%; SD, 35.29). This occurred when either they “smelled something fishy" or the citation looked to be of use to the reviewer for their own research or the reviewer was unsure of their own expertise in the particular area. With very few exceptions, the only abstracts or reviews that were pulled were those that have direct web links to the manuscript being reviewed or those that can be accessed through Medline.

Reviewers indicated that they seldom discussed manuscripts with anyone else. This, they said, is largely due to the instructions they receive from the journal regarding the importance of confidentiality (Dalton 2001, 102). However, only 42.6% (29/68) of reviewers said that they *never* have such discussions. The most common reasons given for discussing a manuscript with a colleague, other reviewer or editor were: the clinical or scientific content is outside reviewer’s area of expertise, or there is a methodological or statistical problem that the reviewer does not feel competent to handle, or the research is exceptional, that is, it is extremely good, bad or provides a good teaching point.

Though the described mechanics of reviewing were remarkably uniform, 85.3% (58/68) of respondents thought their reviews had changed significantly over time. The most common changes described were that reviews were shorter (31% [18/58]) with less copy-editing (24.1% [14/58]) and that there was more focus on larger issues (24.1% [14/58]), clinical significance (19% [11/58]), methods (37.9% [22/58]), and statistics (17.2% [10/58]). An equally common perception was that the quality of their reviews had improved over time (25.9% [15/58]).

It needs to be kept in mind when assessing interview material that what respondents say they do does not of necessity correspond to what they actually do. Further examination of the actual mechanics the review process would require a series

of detailed observational ethnographies to determine how close the correspondence is between statement and activity.

III. Peer Reviewer Attitudes towards Manuscripts, Reviews, Reviewers and the Peer Review Process

Respondents were questioned on and encouraged to discuss what they believed to be the desirable attributes of:

1) research manuscripts,

2) manuscript reviews,

3) peer reviewers, and

4) the peer review process.

TABLES: "Percentage responded" is the number of subjects responding to each question, not percentages of the total number of responses (multiple responses to questions were allowed).

Table III.1) The Research Manuscript What makes a strong article?	percentage responded	total responses
Sound study design / methodological rigor	86.8%	59/68
Clinically relevant	66.2%	45/68
Well-written	42.6%	29/68
Original	22.1%	15/68

Table III.2) The Manuscript Review What makes a good review?	percentage responded	total responses
Provides authors with ways of improving the manuscript	73.5%	50/68
Identifies major weaknesses (and strengths) in the research	57.4%	39/68
Presents its criticism in a positive way	44.1%	30/68
Gives the editor indication of whether to publish or not to publish	33.8%	23/68
Is thorough, showing evidence of time and effort having been spent	27.9%	19/68
Fair	11.8%	8/68
Which is more important: evaluation or instruction?		

Instruction	51.6%	32/62
Evaluation	33.9%	21/62
Both	14.5%	9/62
Change in reviews over time:		
Reviews have changed over time	85.3%	(58/68)
More focus on methods	37.9%	(22/58)
Review is shorter	31%	(18/58)
Quality has improved	25.9%	(15/58)
Less copy-editing	24.1%	(14/58)
More focus on larger issues	24.1%	(14/58)
More focus on clinical significance	19%	(11/58)
More focus on statistics	17.2%	(10/58)

Table III.3) The Peer Reviewer What makes a competent peer reviewer?"	percentage responded	total responses
Knowledge		
Knowledge of the subject area	54.4%	37/68
Methodological expertise	30.9%	21/68
Statistical expertise	10%	7/68
Skills		
Thoroughness	45.6%	31/68
Critical skills	20.6%	14/68
Promptness	16.2%	11/68
Good communication skills	10%	7/68
Experience		
Research experience	41.2%	28/68
Clinical experience	17.6%	12/68
Review experience	9%	6/68
Ethics		
Fair	27.9%	19/68
Supportive	19%	13/68
Education and Training	7.3%	5/68

Table III.4) The Peer Review Process	percentage responded	total responses
Does the peer review process accomplish what it should?		
Yes	86.8%	59/68
Sometimes	10.3%	7/68
No	2.9%	2/68
Goals of peer review process		
Keeps poor research out	58.8%	40/68

(evaluative function)		
Helps improve individual manuscript (instructional function)	39.7%	27/68
Defects of peer review process		
Bias		
Bias of individual reviewers	36.8%	25/68
Selection bias in choosing reviewers	45.6%	31/68
Tone: the "mean-spirited review"	21%	14/68
Research Validity:		
Poor research gets published	14.7%	10/68
Good research is discouraged	23.5%	16/68
Time		
Time consumed doing the review	20.6%	14/68
Timeliness in publication	22.1%	15/68
Blinding		
Continue blinding	73.5%	50/68
Discontinue blinding	14.7%	10/68
Doesn't matter	11.8%	8/68
Improvements to peer review process		
Increase recognition of value of peer review	50%	34/68
Increase feedback from editors	42.6%	29/68
Improve reviewer selection process	25%	17/68
Further training	13.2%	9/68

Table III.5) Motivation and Reward	percentage responded	total responses
Motivations for doing peer review		
Duty: (a and/or b)	70.6%	48/68
a. Responsibility to specialty	70.6%	48/68
b. Reciprocal responsibility to other researchers	22.1%	15/68
Education: Improve knowledge and skills	64.7%	44/68
Career Advancement	32.4%	22/68
Satisfaction: Personal satisfaction / It's an honor	23.5%	16/68
Benefits/rewards of peer review		
Duty: Satisfaction at paying back a debt	32.4%	22/68
Education: Improves knowledge and skills	60.3%	41/68
Career Advancement	27.9%	19/68
Satisfaction: Good feeling	26.5%	18/68
Are the rewards of reviewing are sufficient?		
Yes	82.5%	56/68

No	17.6%	12/68
How might reviewers be rewarded?		
Satisfaction:		
Tokens of appreciation	44.1%	30/68
Thank you notes / publish name in journal	44.1%	30/68
Awards	32.4%	22/68
Career Advancement		
Increased weighting in performance evaluations	42.6%	29/68
Education		
Training courses	20.6%	14/68
Should reviewers be paid?		
No	57.4%	39/68
Yes	42.6%	29/68

DISCUSSION

Whether entirely realized in practice or not, the above tables demonstrate what peer reviewers believe the process of peer review to be and constitute a normative model of that process

Though a number of studies have demonstrated that the effect of peer review on the quality of published research is questionable, 86.8% [59/68] of respondents thought that the peer review process accomplished its goals of validating scientific fact and advancing medical therapeutics (Cole 1981, 881; Goodman 1994, 11; Jefferson 2002, 2785; Pierie 1996, 1430; Rothwell 2000, 1964). Suggestions as to how to improve the peer review process were limited. Though there were some interesting individual observations (e.g., on-line link for journal readers between manuscripts and reviews of the manuscripts; declaration of interest by reviewers, etc.), the only interventions for improvement mentioned more than sporadically were: (Table III.4)

a) Increase the value of peer review in academic performance evaluations (50% [34/68] of respondents).

b) Increase feedback from editors (42.6% [29/68]).

c) Improve reviewer selection process (25% [17/68]).

d) More training courses (13.2% [9/68]).

82.5% of respondents found the rewards offered by the process sufficient for their purposes. Further, the ease with which the categories of response concerning rewards and motivations map onto each other would seem to indicate that, at least for this set of

reviewers, the process functions in a way that meets their expectations (Table III.5). It might be argued that the rather limited list of suggestions for improvement of the peer review process and the close mapping of motivations and rewards confirms reviewers' stated overall satisfaction with peer review.

Tensions in the Process of Peer Review

Though, for the most part, this normative model conforms to conventional ideas about peer review, and the preceding data appears to suggest agreement by reviewers with those conventions, reviewers did express major concerns about the process. The anxieties felt by reviewers were not based on disagreement about goals, functions, criteria, methods, etc., but rather on agreement over tensions inherent in the process that they see as causing contradictions in peer review's operations and outcomes. An identifiable set of interconnected tensions was defined by respondents in relation to:

a) The function of peer review: evaluation and instruction.

b) The goals of peer review: prevent the dissemination of poor research and improve the quality of published research.

c) The structure of peer review: a triage system of checks and balances coupled with a collaborative dialogue of improvement.

d) The motivations and rewards of participation in peer review: duties to researcher, discipline and self.

1) Functions of the Peer Reviewer: Evaluation and Instruction

Respondents cited two main functions to peer review: evaluation and instruction. As an evaluative exercise, peer review was seen as a process of triage—selecting submissions for publication by filtering out incorrect or inadequate work, that is, preventing poor research from being published. This evaluative, triage or censoring function was constructed as expert judgment exercised in service to scientific progress, therapeutic efficacy and disciplinary validity.

> *"Purpose of a review is to provide for the journal's readers the best possible information. Review is a screening tool. Want the regular emergency doc to be able to read the manuscript once and walk away with useful information" (17).*
> *"It identifies manuscripts that advance the specialty" (52).*
> *"I want something to say about what gets published in my discipline" (69) .*

In terms of an instructional function, respondents believed that a good review should also improve the accuracy, clarity and utility of published research. In providing the author/researcher with concrete suggestions for improving the manuscript, the review was thought not only to help improve that manuscript but also to help improve the skills of the individual researcher.

> *"There is something valuable in every paper so I always try to mention the strengths – but pointing out the weaknesses is the most important thing in improving the paper….Give suggestions for correction of errors and solutions for weaknesses" (47).*
> *"Even if the paper doesn't get published, the review should help improve the skills of the researcher" (87).*

The balance between instruction and evaluation was shaky. When asked what makes a good review, 73.5% (50/68) of respondents held that it was one that provides authors with ways of improving the manuscript (instructional function), and 33.8% (23/68) said it was one that gives the editor indication of whether to publish or not to publish (evaluative function) (Table III.2). When directly asked to rank instructional and evaluative functions, 51.6% (32/62) ranked instruction as more important, 33.9% (21/62) ranked evaluation as more important, and 14.5% (9/62) refused to rank one above the other (Table III.2). However, when asked about the goals of the peer review process, 39.7% (27/68) of respondents stated that it should improve the individual study (instructional function) and 58.8% stated that it should insure scientific validity by keeping poor research out of the literature (evaluative function) (Table III.4). As the overwhelming majority of respondents insisted that evaluation and instruction were complementary and that both were necessary, there is perhaps good reason not too make too much of these variations in ranking.

> *"The manuscript review has multiple mission : Trying to help the author but also to determine if the manuscript would appeal to that particular journal and if it is any good" (104).*
> *"Generally, a competent review should be able to come to a decision over whether the paper is basically valid and relevant enough to publish…. It's no good just saying that something is good or bad, but you should be able to say why so the author can correct it" (4).*

Nevertheless, though respondents were well aware that evaluation and instruction were complementary and not mutually exclusive, in discussions, particularly involving the "tone" of reviews, there was concern over the difficulties involved in balancing the roles

of instructor and evaluator, mentor and censor, peer and reviewer (Arrington 1995, 250; Berkenkotter 1995, 248; Fontaine 1995, 261).

> *"I used to be more concerned with criticizing the manuscript, picking out reasons not to publish, now I try more to help the author improve the manuscript" (88). "Triage and instruction are important. My job is the first: my duty is the second. My job is to filter what gets published. My duty is to improve the world of research both published and performed" (60).*

It has been noted that this tension between evaluation and instruction is rooted in peer review's historical development (Biagioli 2002, 20). Peer review evolved out of, on the one hand, government control and censorship of scientific publication in the 17th century and, on the other, the increasingly collaborative production of scientific knowledge in the 18th (Shapin 1987, 417). The result being that almost right from its inception peer review had the dual functions of preventing the dissemination of poor research and assisting in the production of good research.

2) Goals of Peer Review: Prevent Dissemination Poor Research and Encourage Good Research

Related to respondents' perception of the dual functions of peer review was a similar duality of goals. Despite the belief of the majority of respondents that peer review accomplishes what it should, concerns were expressed that peer review often failed at exercising a high selectivity for poor papers and good sensitivity for important research[1] , i.e., poor research is getting into the literature and important research is being overlooked. Poor research slipping into the literature is a major concern of editors:

> "One trouble is that despite this system, anyone who reads journals widely and critically is forced to realize that 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." (Rennie 2002, 2759)

[1] Good research was primarily valued as improving therapeutic efficacy (disciplinary practice), whereas poor research was seen as a threat to the validity of disciplinary knowledge (disciplinary theory). Though practice and theory are inextricably linked, and good research strengthens disciplinary knowledge and poor research is a threat to patients, that was not, for the most part, how concerns over good and poor research were expressed.

The problem of poor research being published was cited as a defect in the peer review process by 14.7% (10/68) of respondents. However, the complementary concern, that of innovative research being "strangled in its cradle" was cited by 23.5% (16/68). There was a fear that much good research is slipping thorough the cracks because ideas that are unusual or methods that are different are likely to be either misunderstood or undervalued by at least one reviewer on a panel. The review process was seen as being inherently conservative, tending towards the status quo, and thus impeding scientific progress and therapeutic innovation (Finn 1986, 14-16; Fontaine 1995, 260)

> *"I think manuscripts get rejected that are good manuscripts and truly innovative ideas are lost" (06).*
> *"There is an inherent timidness in the process…. stifles avant garde science and takes the joy out of research" (48)*
> *"There is a danger of good work being stifled rather than encouraged" (01).*
> *"When something is really avant-garde or cutting edge it is likely to be turned down by the self anointed elite" (107).*

There is ample evidence in the literature that peer review often falls down on both counts: disseminating poor research by failing to identify major errors and failing to recognize important advances (Baxt 1998, 310; Godlee 1998, 237; Lock 1993, 382; Nylenna 1994, 149; Wager 2001, 258-9). The point here is not to prove or disprove claims about defects of sensitivity and specificity in peer review, but to indicate that this is a tension inherent in the process of peer review of which reviewers themselves are concerned.

Another concern related to the dissemination of important research was the timeliness of the process. 22.1% (15/68) of respondents thought that the process was often untimely for authors and created delays in publication (Table III.4). However, in the case of the particular journal for which these respondents review this was considered less of a problem, particularly since the journal's implementation of electronic submission and review.

3) Structure of the Peer Review Process: Checks-and-Balances and Collaborative Dialogue

A) Checks and Balances: Objectivity

Respondents believed that: reviews should be objective, reviewers should be unbiased, and the process should be fair. These virtues were thought to be insured by three things: the reviewer's personal integrity, the blinding of reviews, and the make-up

of the review panel. Both individual bias (unconscious bias, politics, animosity) and selection bias (arbitrariness or favoritism in regard to how reviewers are selected and review panels constituted) were considered problems by, respectively, 36.8% (25/68) and 45.6% (31/68) of respondents.

Personal bias was seen to arise from disciplinary politics, conflicts of interest and petty jealousies. This type of bias was seen as being dealt with by blinding. As imperfect as most respondents knew blinding to be, nevertheless, the majority of respondents, 73.5% (50/68), maintained that the review process should continue to be blinded (Jefferson 2002, 2785; Justice 1998, 240). Respondents believed that blinding is necessary to prevent provenance from influencing judgment. However, there was also concern that hurting an author's feelings could engender animosity against the reviewer, raising the possibility that blinding may be as much for the reviewers' comfort as for its effect on objectivity.

> *"Anonymous reviews are more objective and you don't have to worry about hurting your friends" (22).*

14.7% (10/68) of reviewers in the present study thought that blinding should be abolished and 11.8% (8/68) thought that blinding made little difference one way or the other to the quality of reviews. The two most common reasons given for discontinuing blinding were:

a) It would make possible more direct dialogue between reviewer and author leading to improvement of the manuscript.

b) It would improve the "tone" of reviews. That is, decrease the incidence of "mean-spirited" reviews.

In a prior article, one of the authors of this paper has pointed out that: "The *Annals* continue to believe there is a benefit to this practice [blinding], if only in that the authors' confidence in the review may be increased. As in conflict of interest, the perception that the process is fair and impartial is probably as important as the actual impartiality itself—that is, the blinding policy sends a message about the journal's priorities and concerns" (Callaham 2002, 313).

Though respondents believed that blinding and the skills and ethics of individual reviewers determine the objectivity of the individual review, they also believed that the fairness of peer review as a process is determined by the corporate nature of the review panel.

> *"Each reviewer has their own strengths – so there is a variety of talents working to improve the quality of the research" (47).*
> *"I think the validity of the review process operates more at the level of…. assembling the right team of reviewers. Each reviewer brings his own view to the article. It is not good to have a team of reviewers with all the same abilities and strengths" (86).*

A balanced review team made up of qualified individuals with a variety of knowledges and skills was thought to increase the probability of producing a fair decision about publication. However, this could be influenced by selection bias. Selection bias was seen to operate at two levels:

a) Reviewer Pool - the selection of researchers to become part of the journal's reviewer pool.

b) Review Panel - the selection of a panel of reviewers to review a particular manuscript.

Concerns over selection bias at these two levels appeared to be paradoxical:

a) Reviewer pool – The major concern was over too much variation in the quality of reviewers in the reviewer pool.

> *"There is a lot of individual variation between reviewers. As an author it's luck of the draw who you get as a reviewer"* (74).
> *"Journals should… be more selective about who they ask to review" (22).*

b) Review panel – The concern was too little variety in the viewpoints of reviewers on the review panel.

> *"The* Annals *seems to use the same reviewers all the time. I suppose because the specialty is so small. I think there's a danger in this, in that it could lead to the only a certain type of research being published" (03).*
> *"Need broader pools of reviewers. Need to go beyond the usual sets of reviewers to increase the richness of the pool" (70).*

Thus, on the one hand, the requirement for a large pool of reviewers was seen as resulting in the recruitment of reviewers with inadequate skills, producing unacceptable variation in the quality of the reviews. On the other hand, there was equal concern that, if too small a group of skilled reviewers were used to make up review panels, reviews would become too uniform with the same type of manuscripts being published and the same type rejected. Unsurprisingly, when respondents were asked how they would improve the peer review process, recommendations for improving reviewer selection involved both narrowing and broadening the pool of reviewers.

Overall, respondents described a system where the fair assessment of a manuscript depends upon the right mix of qualified reviewers, whose different expertises and skills complement each other, presumably canceling out particular ignorances and biases. Though the goal of the process is to come to some kind of consensus on what constitutes valid science, the process maintains its own internal validity through the variety of the reviewers' knowledge and skills, the independence of their judgment (independence from personal bias, editorial pressure and peer influence), and their potential for disagreement (Emanuel 2005, 2559). Peer review was thus seen as a system of checks and balances, requiring a degree of separation of authorial, review and editorial functions, achieved through a combination of blinding and multiple independently generated reviews. The most frequent analogy used by respondents when discussing both the strengths and weaknesses of peer review was "democracy".

"Peer review is like democracy" (62).
"Three heads are better than one" (20).
"Like the American system it's not perfect but it is the best thing we have" (19).

It perhaps should come as small surprise, considering the nationality of most respondents, that peer review is likened to the American style of representative democracy heavily dependent on checks and balances (as opposed to parliamentary democracy, geared more towards negotiated consensus). It has been pointed out by others that the present system of peer review places its faith in a median being established through the balancing of opposing views (Finn 1986, 14).. Such a system is inherently conservative with a tendency to privilege the status quo over radical change—not necessarily a bad thing, either in politics or science. However, concern about the deleterious effects of such conservatism was reflected in respondents' comments about innovative research being stifled and in concerns over communication between reviewers and authors, and between editors and reviewers.

B) Collaborative Dialogue: Feedback

Despite respondents support for the current system of checks and balances and that system's requirement for a degree of separation of authorial, review and editorial functions, many respondents expressed the desire for a more open system of dialogue with increased feedback. Two issues arose repeatedly in the interviews in relation to feedback:

a) condemnation of "mean-spirited" feedback from reviewers to authors and

b) frustration at perceived lack of feedback from editors to reviewers.

These concerns over feedback surfaced in multiple places in the interviews: in discussions of skills' acquisition, training, evaluation, motivations, rewards, defects in the process, improvements to it, and in the relations between editors, reviewers and authors.

Reviewer>Author: Feedback and "Tone"

In discussions about the function of the review, respondents claimed that they give feedback to authors to improve the submitted manuscript and to assist the researcher in improving his or her research skills. A desire for, but problems with, increased reviewer>author dialogue was made evident in attitudes towards blinding. While most respondents supported blinding and shied away from constructing reviewer>author feedback as a direct encounter, most also believed that an important function of peer review is to provide the author/researcher with the necessary information, tools and encouragement to successfully get their work published. As already discussed, a minority of respondents (14.7% [10/68]) did recommend the removal of blinding to create a more open system:

> *"It's a closed process that allows for only limited back and forth dialogue. Sometimes my criticism is wrong because I've misunderstood or misread something, and the author should have a chance to correct me… It would be useful to give both the reviewer and the author opportunity to contact each other." (01).*
> *"It [blinding] stifles dialogue. There are models of open review where things can be discussed by reviewers and authors and a lot of insight gained" (01).*

The big concern related to reviewer>author feedback was the issue of "tone". This concern has been noted by other authors (Fontaine 1995, 259; Siegelman 1991, 637).

> "What emerged from most writers' descriptions of their experiences is a perceived antagonistic relationship between themselves and reviewers created by the tone and content of reviewers' comments…" (Fontaine 1995, 259).

When asked about the characteristics of a good review, 44.1% (30/68) of respondents thought that it was important for criticism to be presented in a positive manner.

> *"[Reviews should] not be critical just to be critical, but critical in order to help strengthen the work…. Some people are assassins and some people are saints and it helps to bring the two ends together" (45).*

21% (14/68) of respondents cited the "mean-spirited" review as one of the major defects of the peer review process.

> *"Some reviewers have a real mean streak in them….I was getting reviews back that were mean-spirited and not helpful in terms of fixing the manuscript. Often the reviewers simply did not understand the paper. It motivated me to improve my own reviewing skills" (93).*

19% (13/68) of respondents listed "supportiveness" as an important attribute of a competent peer reviewer. This concern over reviews that were mean-spirited, arbitrary or hypercritical was raised in other key areas in the interview, as when respondents discussed how their reviews had changed over time and ways of improving the peer review process. A not uncommon narrative of reviewer development over time involved a progression in reviews from a commentary that was necessarily vague due to reviewer inexperience, through a period of hyper-critical comment as skills were acquired, on to a more constructive critique focused on assisting the author.

> *"When I first started I didn't have a focused method. I just gave vague impressions. Then I evolved into a devil's advocate type role – really picking apart and criticizing manuscripts. But now I try to be more supportive – I try to make the manuscript as good as possible. It's like working with a colleague to make the manuscript better…. My overall approach is no longer just to pick the study apart and show what is wrong – I've moved beyond that – to a more productive attitude – I'm no less critical but I do it in a different way." (93).*

Considering the almost universal condemnation of the mean-spirited reviewer, one wonders who it is that is writing these negative reviews? Reviewers not contacted by this study? Or is what appears to be a detailed and conscientious critique when given as a reviewer read as mean-spirited when received as an author? Concern about the tone of reviewer>author feedback was tied to the peer reviewer's ambivalence over the dual functions of instruction and evaluation. This was especially evident in reviewers' desire to motivate researchers and encourage good research.

> *"Certain journals include the derogatory comments in the letter to the author – authors need to believe in themselves and this kind of behavior doesn't help" (47).*
> *"Sometimes the reviews are so intimidating they discourage good authors from getting published" (75).*
> *"I end up struggling to find faults… that really aren't very helpful to the author and may even frustrate him" (05).*

<u>Reviewer>Reviewer: Feedback as Dialogue</u>

Respondents claimed that access to other reviewers' reviews was the most effective means of "on the job" skills' acquisition. The importance of reviewer-reviewer

dialogue was emphasized in comments about the utility of and satisfaction gained from reading the other reviewers' reviews.

> *"Reviewing is quite isolated. Often don't know where one stands. Not a lot of interaction – so it is useful to see what others are doing" (43).*
> *"The more you are involved in a discussion and able assess your own reviews against other people's comments, the better. It allows you access to other people's thought processes" (15).*

What most reviewers look for in the other reviewers' comments is what they themselves have missed. It is a way for them to judge *the quality of their review* and so improve their own review skills. The value placed on peer dialogue and its role in skill improvement was also evident in reviewers' comments about participation in journal clubs, grants and IRB committees. The "crossover" activities that respondents participated in, which they cited as improving manuscript review skills, were: journal club (91.2% [62/68]), grants' review and IRB committees (63.2% [43/68]), and teaching critical appraisal, study design or epidemiology (50% [34/68]). Although journal club was the most frequent activity, participating in grants' review committees was considered the most useful in sharpening critical skills used in manuscript review. Respondents placed emphasis on the opportunities that the committee format opened up for critical dialogue between peers.

> *"Grant review committees involve interactional feedback. It's a collaborative process which helps improve skills." (56).*

They claimed that their ability to engage in dialogue produced what they thought were better reviews than individual critique may have done.

> *"Grant review and research committees are very useful as everyone reviews the same material and you get to discuss it with a group of people" (82).*

The use of manuscripts as teaching tools for residents and fellows was also seen as another means of establishing a dialogue between peers around scientific work.

> *"'I think the most useful [way of teaching review skills] is to pull apart selected manuscripts as a group with the goal of teaching specific points about what to look at" (88).*

<u>Editor>Reviewer: Feedback as Motivation</u>

The most frequent form of editorial feedback occurred in the form of the editor's letter to the author which is shared with the reviewers. Reasons for looking at the other reviewers' comments and the editor's decision letter, though overlapping, were somewhat different. The editor's letter was looked at to see whether the decision to publish or not publish was confirmed, that is, to see whether the reviewer's own judgment had been validated, i.e., it is mostly about agreement on *the quality of the manuscript*. However, useful as this is, its helpfulness in improving skills was considered limited.

> *"I'm interested in knowing what the final disposition of the manuscript is and whether the editor agrees with me or not, but I can't say I learn much from the editor's letter to the author. I would prefer a direct assessment of my review by the editor" (01).*
> *'[Scores] are not that helpful. What would be more helpful would be specific comments by the editor about the review. Constructive criticism. What I did good on, what I did bad on' (28).*

Direct editorial feedback to reviewers about their reviews was, with few exceptions, perceived as being insufficient. Though, respondents' comments about lack of feedback were framed in the language of pedagogy:

> *"I would like more feedback from the editors, so that I know if I'm doing a decent job or not. Telling me what I should be focusing on…. Everyone wants to do the best job they can" (05).*
> *"There should be some kind of mentoring process….. The CD is helpful, but it doesn't close the information loop. You need to discuss your reviews with someone and get guidance" (37).*

there was a distinct emotional quality to their complaints:

> *"I have asked for feedback from editors – I feel that I'm on my own" (93).*
> *"I would love to have insight from editors as to what they consider a good review…. I don't know because no one has ever told me" (19).*
> *"Most of us want to do a good job and more guidance would be welcome and encouraging" (74).*
> *"Reviewers should be rewarded by getting more feedback" (43)*

It is, perhaps, not surprising that those who participate in peer review are invested in the process not only practically, intellectually and professionally; but also personally and emotionally—that they are as concerned about motivation, both as researchers and reviewers, as they are about instruction.

When asked how they would improve the peer review process, 42.6% [29/68] of respondents suggested more feedback. Despite peer reviewers' strongly expressed belief in the importance of feedback, previous studies indicate that editorial feedback to reviewers has little, if any, effect on performance (Callaham 2002, 2781). Whether reviewers' claims about lack of feedback are justified or not, and whether feedback is an effective pedagogical tool or not, the strength of reviewers' faith in the efficacy of feedback and their often emotionally phrased concerns over deficits in its quantity and quality would appear to indicate that as a disciplinary dialogue it is seen as somehow lacking. Respondents indicated that such dialogue carries an important affective charge, and as such is as motivational as it is instructional. Similar issues around feedback and the desire for what has been labeled "dialogic collaboration" have been noted in studies of peer review in journals of composition studies (Fontaine 1995, 259; Hunter 1995, 265).

4) Motivations: Self-Improvement and Duty

Respondents' declared motivations and perceptions of rewards coincided to a large degree and were easily mapped onto four major categories—skills acquisition, career advancement, duty and personal satisfaction—which in turn reflected two general concepts: self-improvement (skills acquisition and career advancement) and duty (fulfillment of duty to peers and to the discipline) (Table III.5).

Self-Improvement:

Skills Acquisition and Training: Skills acquisition through exposure to current clinical knowledge and through exercise of critical skills was the most frequently cited benefit (60.3% [41/68]) of doing peer review and the second most frequently cited motivation (64.7% [44/68]) for participating in the process. Respondents argued that an understanding of the review process and acquisition of review skills were essential to improving their own research skills. Review skills were described as being acquired in a number of ways: attendance at formal training workshops, crossover activities, "on the job" experience, and feedback.

"On the job" experience: Respondents asserted that review skills were primarily acquired "on the job". Three modes of experience were deemed necessary to produce a competent peer reviewer: experience doing research, clinical experience, and experience doing reviews. Experience doing research (41.2% [28/68]) was mentioned

twice as often as clinical experience (17.6% [12/68]), and clinical experience twice as often as review experience (8.8% [6/68]). Experience doing research encompassed not just the active production of research by the reviewer themselves but experience of the review process by the reviewer as an author. That is, respondents felt it was important that they knew what it was like to get back reviews of their own work, which then in turn shape their own practice of review. This, more than anything else, was emphasized as being the basis of the status of "peer" and was seen as productive of a valued empathetic identification of the reviewer with the author:

> *"They [reviewers] should be someone that does research themselves – whose been on the other end and received reviews and had to modify their work. It helps you see how the whole system is put together" (38).*
> "*You can't be a good reviewer unless you've been an author and had experience of success and failure" (47).*

Training programs: 80.9% (55/68) of reviewers had some formal training in peer review in the form of in-house workshops, seminars, or as part of a residency, fellowship or advanced degree. Though 80% (44/55) of those who had done in-house training courses said they found them useful, they qualified this by stipulating that the usefulness of training courses was in introducing junior reviewers to the process of review and that such workshops were of decreasing importance once experience had been gained.

> *"The workshop was useful at the time, early on when I began reviewing… it would probably not be so useful now" (83).*

Prior studies have shown that half-day workshops improve reviewer self-ratings (that is, reviewers believed such workshops to improve their performance), but do not improve actual performance of reviews, or any improvements are small and fade over time (Callaham 1998, 318; Schroter 2004, 673). In a similar vein, a meta-analysis of journal-club formats and critical appraisal courses for residents concluded there was little evidence supporting their efficacy (Norman 1998, 177). Attitudes of reviewers in this study towards training courses reflect the above. Although reviewers experience such training as useful, they appear to give it low priority, with only 7.3% (5/68) of respondents citing training as contributing to the competence of a reviewer, 13.2% (9/68) suggesting that more courses for reviewers might be a way of improving the peer review process, and 20.6% (14/68) suggesting training courses as a form of reward. Although these last two numbers are not insignificant, they rank below other suggestions for improvement

and reward. However, it was pointed out by a number of respondents that the in-house workshops, besides communicating information, also functioned as occasions for dialogue with other reviewers, editors and “leaders in the field”.

> *“They should re-institute the training sessions…. They exposed reviewers to leaders in the field…, helped communicate attitudes, and helped improve morale” (12).*
> *“At our retreats we will discuss difficult manuscripts and it’s very useful to get in a dialogue with the editors” (39).*

Workshops were valued not merely as a vehicles for skills acquisition but as motivational: morale boosters and a means of rewarding reviewers for doing reviews. A significantly larger proportion of reviewers who were also editors (54.1%) participated in in-house training than reviewers who were not editors (17.9%). The positive correlation between editorial position and attendance at workshops raises a question about whether attendance at such training programs may have as much to do with being evidence of or fostering commitment to the process as with any direct influence on skills acquisition.

Career Advancement: 50% (34/68) of reviewers cited advancing their careers, increasing their prestige, or networking within the discipline as motivation for doing peer review. The issue of career advancement tied in with concerns respondents had about time spent reviewing. Though 79.4% (54/68) of reviewers stated that reviewing did not take up an inordinate amount of their time (mostly because of their ability to limit the number of reviews they agree to do), 20.6% (14/68) commented that the process was time consuming in relation to the recognition it receives from academic departments. The low weighting of peer review in academic departments' performance assessments (upon which career advancement is based) was a concern of 50% (34/68) of the respondents. They pointed out that other academic activities give a "better return" for time invested.

> *"More weight should be given to doing reviews by academic departments when they are promoting people…. As a division chief, I have difficulty in getting faculty motivated to review. They say, ‘I get to put a line on my CV, but it doesn’t really help me.’ There needs to be more academic recognition, more points given in the advancement process for doing reviewing" (12).*

42.6% (29/68) thought reviewers should be rewarded by increased weighting of peer review in performance evaluations. Respondents argued that peer review merits this not merely as a service to the profession but as an effective mode of professional development. Peer review was seen as *‘improving both knowledge and skills’ (96)* by

keeping the physician on top of recent clinical and scientific developments and by improving research and critical skills. Furthermore, peer review was also thought to improve teaching skills, as the critical skills learned in the practice of peer review are then passed on to interns, residents and colleagues.

> *"Every medical journal in the US should make a statement of position to medical school deans about the importance of peer review to the discipline, the individual and the institution" (45).*

Other authors have made convincing arguments for review as a form of scholarship influential within the discipline and thus worthy of serious recognition by academic departments (Emanuel 2005, 2560; Hesse 1995, 254).

Payment: Payment for reviewing was *not* brought up spontaneously by any respondents, but, when introduced into the discussion by the interviewer, 57.4% (39/68) were against payment and 42.6% (29/68) thought it might possibly be useful. However, even within that group who thought payment might be useful, support for payment was lukewarm. When it was considered, payment was thought to be best offered as an honorarium rather than an attempt at reimbursement. The most frequent reason given for instituting payment of reviewers was that it might improve the quality of reviews by encouraging reviewers to spend more time doing a more conscientious job. An interesting variation was brought up by a couple of respondents who construed payment less as a reward for the quality of the product produced than as a mark of the value of the labor that goes into it.

> *"In the new climate of 'corporate' academics, there is pressure by the institutions on individuals and departments, that if it isn't bringing in money then they don't want to know about it. There is less appreciation of the importance of academic duties… Payment gives 'real value' to the work and product."* (34).

It was argued that as academic institutions are forced to become more and more mercenary, the ability to bring in money becomes a mark of the value of the individual researcher. The implication is that paying reviewers (even an honorarium, where payment is but a symbolic mark of value) might give more weight to peer review in the performance assessments done by academic departments. A common reason given against payment was that it would be financially untenable for the journal, increasing the costs of production unreasonably. The ethical issue was also raised; payment was seen as possibly creating conflicts of interest. Concern was expressed that payment might

detrimentally alter the make-up of the pool of reviewers, turning it into a small group of "professionals" resulting in reviews becoming too uniform. A small pool of paid, expert peer reviewers was seen as little more than an expanded editorial board, where individual biases might hold sway and the checks and balances that are seen as the strength of the peer review process lost.

Duty

Duty Required: To the Discipline and to the Researcher: "Duty" was the most frequently cited motivation (70.6% [48/68]) for being involved in peer review. Duty was construed as operating at a number of overlapping levels: to science, to the discipline, to researchers and to the patient. Respondents conceived of peer review as a crucial component of scientific progress, most often interpreted in practical terms as improved therapeutic efficacy. The reviewer's responsibility to therapeutic efficacy was seen not merely as a technical but as an ethical imperative, i.e., what is best for the patient.

> "*Peer review shapes the direction of research…. directing it along more fruitful lines… it's part of my responsibility—to insure that the academic literature is rigorous and current. I owe it to doctors and patients that information out there is accurate" (42).*

The activity of peer review was seen as a crucial point in the transformation of individual research into valid disciplinary knowledge, which, in turn, was seen as necessary to the viability and status of the discipline.

> *"It helps the specialty stay valid in the eyes of other specialties" (05)*

The operation of disciplinary knowledge was perceived as relying upon a system of reciprocity whereby reviewers see themselves as logically obligated to participate in the process, providing for others the necessary evaluative service of review that they themselves required to get their own work validated and published.

> *"I see it as an exchange. As I submit papers and want them to be peer-reviewed and published, so I think that I in turn have a duty to review other people's papers" (08).*

Reciprocity was also seen as an ethical imperative, driven by an empathetic identification on the part of the peer reviewer with the author.

> *"Every paper is an opportunity to help a colleague" (14).*

> *"Often I will do extremely meticulous reviews of bad articles because I see it as way of paying back people who spent time on my own error-filled work" (28). "Someone gave their time to help me and I should do likewise" (30).*

Though respondents indicated a sense of loyalty to the journal, which might be rolled into duty to the discipline, many respondents also expressed a personal loyalty to the individual editors with which they worked. The respect in which the editor-in-chief and the section-editors were held was given by a number of reviewers as the reason why they participate in the peer review process.

> *"This editor has really spear-headed the approach to EBM. Loyalty is an important part of building up the informal system of give and take that helps keep the system operating" (99)*.

In light of the above, the sense of duty, an important motivation for the majority of peer reviewers, appears to be based not only on the logical imperative for systemic reciprocity but also on an ethical obligation to one's peers driven by empathetic identification with researchers and personal loyalty towards editors.

Duty Fulfilled: Personal Satisfaction: 44.1% (30/68) of respondents reported "satisfaction", "enjoyment", and "a good feeling" as motivations for participating in peer review. These respondents claimed that they experienced satisfaction in helping others and enjoyed the thanks expressed by the journal and editors. The form of such recognition appeared to matter less than its fact. The annual listing of reviewers' names in the journal and letters of appreciation from the editor were particularly well received. Respondents who had received top consultant awards showed great pride in them.

> *"Being listed as one of the 'Top 50 Consultants' was a big deal for me. I really felt that it validated what I try to do and let me know that I'm doing a good job" (05).*

Again, as in their construction of "duty", respondents emphasized the importance of the affective factors.

> *"It's a challenge. I enjoy doing it" (107).*
> *"It's a question of fighting the good fight" (96).*
> *"The value is in pride of accomplishment…. Makes me feel good about what I do" (33).*
> *"Its good to know that you help bring quality work out and help make authors better researchers" (22).*

The peer reviewer's desire for the improvement of his or her own skills, or his or her own professional position, is by no means directly in conflict with the reviewer's duties

towards the discipline and towards other researchers. The exercise of the reviewer's duties assists in the construction of disciplinary knowledge and the maintenance of disciplinary validity. Self-improvement, not only through the acquisition of knowledge and skills but through peer interaction, contributes to the development of the individual's disciplinary identity. Nevertheless, the mixed set of duties that reviewers' feel they have to researchers, editors, patients, the journal, and the discipline appear to be taken very seriously and many of the respondents expressed anxieties about their ability to get the balance right.

CONCLUSION

A Normative Model of Peer Review

This study describes a normative model of peer review based on the beliefs held by peer reviewers about the process. Reviewers believe that the peer review process should be:

1) Effective: Peer review should promote therapeutic progress by improving the quality of published researched.

2) Valid: The process should accurately assess research to insure that poor research is kept out of the literature, and important research gets published,

3) Fair: It should objectively assess the quality of the research.

4) Efficient: It should be timely for the author and time well-spent for the reviewer.

To achieve those goals, reviewers describe the desirable characteristics of manuscript, review, reviewer, and process:

The manuscript: The research evaluated should have sound study design, methodological rigor, be clinically relevant, well-written and original.

The review: The review should be constructed as both evaluative (provides the editor with direction about what should and should not be published) and instructive (provides the author with useful advice for improving the manuscript). Reviewers' primary concerns about reviews are that they identify major weaknesses in the manuscript and present criticism in a useful and constructive manner. The reported mechanics and methods of reviewing showed little variation from respondent to respondent, with the structure of the review paralleling the structure of the manuscript. Though the described mechanics of reviewing were remarkably uniform, the majority of respondents believed their

reviews had changed significantly over time, with the most common changes being that reviews were shorter with less copy-editing and with greater focus on larger issues, clinical significance, methods and statistics.

The reviewer: The ideal peer reviewer should be knowledgeable, skilled, fair and constructive. Unbiased evaluation and constructive instruction were thought to be secured through the selection and education of a pool of quality reviewers. Respondents believed that reviewers should be selected on the basis of the breadth and depth of their knowledge, practice in doing research and evidence of critical skills. Review skills were believed to be improved through experience and training, but most especially through feedback from editors. Respondents stated they are motivated by a combination of a sense of duty, a desire for the acquisition of knowledge and skills, career advancement and personal satisfaction.

The peer review process: Reviewers believe that the validity of the peer review process (and by extension the knowledge it sanctions and the discipline it represents) depends upon critical expertise and objectivity. These are insured by:

1) Skills: The critical skills of the reviewers being improved primarily through participation in peer review and editorial feedback.

2) System Structure: The structure of the review process is conceived of as a system of checks and balances. Objectivity is achieved through the selection of a skilled reviewer pool and the constitution of a diverse review panel. Blinding is seen as structural means of decreasing bias and insuring objectivity.

3) Duty: The reviewer is believed to be motivated by:

a) An ethical duty based on empathetic identification of the reviewer with the researcher.

b) A practical duty to the discipline based on the system's requirement for reciprocity.

The majority of respondents believe that the peer review process does, for the most part, accomplish what it sets out to do. The close agreement between reviewers' declared motivations and perceptions of rewards appears to indicate that the process meets their expectations.

Though this normative model conforms to conventional ideas about peer review, reviewers expressed concerns about the process. The anxieties expressed by

reviewers were not based on disagreement about goals, functions, criteria, methods, etc., but rather on agreement over tensions inherent in the process causing difficulties in peer review's operations and outcomes. In the actual practice of peer review, the mandate to both evaluate and instruct, the goals of identifying poor research and improving good research, and the attempt to participate in a collaborative dialogue within a triage system of checks and balances, requires that a balance be struck between the needs of the discipline and the needs of the individual researcher and reviewer. In their perception of the role of the peer reviewer as both censor and mentor, though respondents cited evaluation of the manuscript (selecting submissions for publication by filtering out incorrect or inadequate work) as the primary goal of the process, instruction of the researcher (improving the accuracy, clarity and utility of published research) was considered by a majority of reviewers as the more important practice. Though there was recognition that the goal of the review process is to prevent poor research from being published, there was greater concern that the process results in innovative research being “strangled in its cradle”. Though reviewers believe that the quality of the review is determined primarily by the skills of the individual reviewer, they maintain that the validity of the process is determined by the corporate nature of the review panel acting as a system of checks and balances. Though perceiving this system of checks and balances as requiring a degree of separation of authorial, review and editorial functions; reviewers, at the same time, expressed the desire for a more open system of feedback leading to collaborative dialogue. Though the peer status of the reviewer, believed by reviewers to be the prime source of their validity as reviewers, is seen as being based on a shared set of specialized skills and knowledge, it is also believed to require empathetic identification with the author. An empathetic identification with the author was seen as being balanced against a duty to the discipline, determined by the system's need for objective reciprocity in the work of review.

The tensions found in the peer review process, sometimes seen as barriers to its effective operation, are less defects in the process than definitive of the concept of peer review itself and thus necessary to its operation. While at a practical level peer review operates as a triage exercise, it is, at the same time, on a social level, a mode of disciplinary dialogue between peers: important not only to the maintenance of an effective knowledge base and thus disciplinary validity, but also, through its effect on researchers and reviewers, important in the construction of disciplinary identity. Peer review defines not only what and whom a peer is, but it is, at least partly, through

participation in the process, that the researcher/reviewer develops the status of "peer". Peer review's practical and social operations are not antithetical to each other but rather are inherent in the hybrid concept of the peer reviewer, where one's status as a peer makes possible one's activity as reviewer.

BIBLIOGRAPHY

Arrington P. Some thoughts on changing the review process for academic journals: a personal exploration. *Rhetoric Review*. 1995;13(2):249-53.

Baxt WG, Waeckerle JF, Berlin JA, et al. Who reviews the reviewers? Feasibility of using a fictitious manuscript to evaluate peer reviewer performance. *Ann Emerg Med.* 1998;32:310-317.

Bazarian JJ, Davis CO, Spillane LL, et al. Teaching emergency medicine residents evidence-based critical appraisal skills: a controlled trial. *Ann Emerg Med.* 1999;34:148-154.

Berkenkotter C. The power and perils of peer review. *Rhetoric Review*. 1995;13(2):245-8.

Biagioli M. From Book Censorship to Academic Peer Review. *Emergences*. 2002;12(1):11-45.

Callaham ML. Training programs for Peer Reviewers. *CBE Views*. (Council of Biology Editors) 1997;20(6):192-3.

Callaham ML, Wears RL, Waeckerle JF. Effect of attendance at a training session on peer reviewer quality and performance. *Ann Emerg Med.* 1998;32:318-322.

Callaham ML. Research into peer review and scientific publication: Journals look in the mirror. *Ann Emerg Med*. 2002;40:313-316.

Callaham ML, Schriger DL. Effect of structured workshop training on subsequent performance of journal peer reviews. *Ann Emerg Med*. 2002;40:323-328.

Callaham ML, Knopp RK, Gallagher EJ. Effect of written feedback by editors on quality of reviews. *JAMA.* 2002;287:2781-3.

Clark A. *Situational Analysis: Grounded Theory and the Postmodern Turn.* Thousand Oaks: Sage: 2005.

Cole S, Cole R, Simon G. Chance and consensus in peer review. *Science.* 1981;214:881-6.

Emanuel L, Greenland P. Peer Review and Professionalism at the *Archives of Internal Medicine*. *Arch Intern Med*. 2005;165:2559-2560.

Finn C. Strengths (and weaknesses) of peer review. *Educational Researcher.* 1986;15(7):14-16.

Fontaine S. With writers' eyes: perception and change in manuscript review procedures. *Rhetoric Review*. 1995;13(2):259-64.

Fowler FJ Jr. *Improving Survey Questions*. Thousand Oaks: Sage; 1995.

Garfunkel JM, Ulshen MH, Hamrick HJ, et al: Problems identified by secondary review of accepted manuscripts. *JAMA.* 1990;263:1369-1371.

Glaser BG. *Doing grounded theory: issues and discussions*. Mill Valley, Ca.: Sociology Press; 1998.

Godlee, F., Gale, C.R. and Martyn, C.N. Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. *JAMA* 1998:280, 237–40.

Goodman SN, Berlin J, Fletcher SW, Fletcher RH. Manuscript quality before and after peer review and editing at *Annals of Internal Medicine*. *Ann Intern Med.* 1994;121:11-21.

Hesse D. Reader reviews as scholarship: a chaotic view. *Rhetoric Review*. 1995:13(2):254-58.

Hunter S. The case for reviewing as collaboration and response. *Rhetoric Review*. 1995:13(2):265-72.

Ingelfinger FJ. Peer review in biomedical publication. *Am J Med.* 1974;56:686-692.

Jefferson T, Alderson P, Wager E, Davidoff F. Effects of editorial peer review: a systematic review *JAMA.* 2002;287:2784-2786.

Journal Citation Reports 2004, ISI Web of Knowledge, The Thomson Corp., http://portal.isiknowledge.com/portal.cgi/jcr?Init=Yes&SID=A2GmHbGABmgJLI@fnB9, accessed 2/15/06.

Justice AC, Cho MK, Winker MA, et al. Does masking author identity improve peer review quality? A randomized controlled trial. PEER Investigators. *JAMA.* 1998;280:240-242.

Kassirer JP, Campion EW. Peer review: Crude and understudied, but indispensable. *JAMA.* 1994;272:96-97.

Lambert, H. "Anthropology in health research: from qualitative methods to multidisciplinarity," *BMJ*. 325, 2002, 210-13.

Langkamp DL, Pascoe JM, Nelson DB. The effect of a medical journal club on residents' knowledge of clinical epidemiology and biostatistics. *Fam Med.* 1992;24:528-530.

Linzer M, Brouwn J, Frazier L, et al. Impact of a medical journal club on house-staff reading habits, knowledge, and critical appraisal skills. *JAMA.* 1988;260:2537-2541.

Locke S. The grossest failures of peer review. *BMJ.* 1993;307:382.

Norman GR, Shannon SI. Effectiveness of instruction in critical appraisal (evidence-based medicine) skills: a critical appraisal. *CMAJ.* 1998;158:177-181.

Nylenna, M., Riis, P. and Karlsson, Y. Multiple blinded reviews of the same two manuscripts: effects of referee characteristics and publication language. *JAMA.* 1994:272,149–51.

Pierie JP, Walvoort HC, Overbeke AJ. Readers' evaluation of effect of peer review and editing on quality of articles in the *Nederlands Tijdschrift voor Geneeskunde*. *Lancet.* 1996;348:1480-1483.

Rennie D. Freedom and responsibility in medical publication. *JAMA.* 1998;280:300–2.

Rothwell PM, Martyn CN. Reproducibility of peer review in clinical neuroscience: is agreement between reviewers any greater than would be expected by chance alone? *Brain.* 2000;123:1964–9.

Schroter S, Black N, Evans S, Carpenter J, Godlee F, Smith R. Effects of training on quality of peer review: randomized controlled trial. *BMJ.* 2004;328:673-5.

Shapin S. O Henry: Review *The Correspondence of Henry Oldenberg*. *Isis*. 1987;78(3):417-424.

Siegelman SS. Assassins and zealots: Variations in peer review *. Radiology.* 1991;178:637-642.

Silverman D. *Qualitative Research: Theory, Method and Practice*. London: Sage: 1997.

Smith, R. Peer review: a flawed process at the heart of science and journals, *J R Soc Med.* 2006;99:178-182.

Stossel TP. Reviewer status and review quality. Experience of the Journal of Clinical Investigation. *N Engl J Med*. 1985;312(10):658-9.

Strauss A, Corbin J. *Basics of qualitative research: grounded theory procedures and techniques*. Newbury Park: Sage; 1990.

Strayhorn J, McDermott JF Jr, Tanguay P. An intervention to improve the reliability of manuscript reviews for the *Journal of the American Academy of Child and Adolescent Psychiatry*. *Am J Psychiatry.* 1993;150:947-952.

Wager E, Jefferson T. The shortcomings of peer review. *Learned Publishing.* 2001;14:257-263.

INTERVIEW TEMPLATE

16) Length of interview:	
22) How did you become a reviewer?	
	respondent generated categories
	Asked by editor
	Part of Fellowship / resident position on editorial board
	Volunteered
	Asked by another reviewer
	Can't remember
23a) How many years have you been reviewing for Annals?	
23b) How many years have you been reviewing total?	
23b) How many reviews do you do a year for Annals?	
23b) How many reviews do you do a year total?	
23b1) Do you review for any other journals?	
23b2) Which journals?	
24a) Are you on the editorial board for any journals?	
24b1) Which journals?	
24b) How many years have you been an editor?	
25) Why did you agree to become a reviewer?	
	respondent generated categories
	Education and Skills
	improves disciplinary knowledge clinical knowledge / allows me to keep up with advances in medicine
	it improves my research skills
	it improves my writing skills
	it improves my critical skills
	it improves my teaching skills
	Duty
	Responsibility / duty / contribution to / profession / discipline / specialty / science
	Shaping of specialty / Advancement of discipline / Say in what gets published
	Reciprocal responsibility to other researchers
	Responsibility to patients
	Personal Satisfaction
	Interested in learning about the peer review / publication process
	I enjoy doing it
	It's an honor
	Asked to
	Career Advancement
	Career progression: Improves CV and useful for promotion / part of academic career
	Networking - way of getting connected to leaders in the filed
31a) What other forms of critical appraisal do you participate in?	

	respondent generated categories
	journal club
	grant / research / scientific review committee
	EBM
	conference abstract review
	IRB committee
	clinical / practice / policy guideline committees
	editing textbooks
	grant writing
	others
31a1) What do you fins are the most useful forms of critical appraisal?	
	respondent generated categories
	grant/review committee
	journal club
	IRB committee
	editing textbooks
	conference abstract review
	practice guideline committees
	EBM
32) On a scale of 1 to 5 with 1 being inadequate and 5 being exceptional, where do you rate your review skills?	
	1 – inadequate
	2 – poor
	3 – acceptable and useful
	4 – excellent
	5 – exceptional
33) What formal training have you had in critical literature review?	
	respondent generated categories
	Annals workshops
	Annals CD ROM
	SAEM seminars
	Advance degrees or training
	Classes on grant writing / grant review / on EBM / study design
	Mentoring
	Self Study: Articles on critical appraisal / EBM
	None at all
34a) Do you teach critical appraisal, study design or epidemiology?	
41) Could you take me through, in a step by step fashion, how you go about conducting and writing a review?	
51a) How long after receiving the manuscript do you start the review?	
41a) How many times do you read a manuscript?	
51b) In how many sittings?	
51c) How long does it take you?	
52a) Where do you do your reviews?	
	respondent generated categories
	Home
	Work
	Travel

	All three- home/work/travel
53a) What percent of the time do you discuss the manuscript with someone else?	
53b) What triggers discussion with someone else?	
	respondent generated categories
	Content
	Content is outside reviewer's area of expertise
	- Methodological or statistical problem
	Quality
	Paper is extremely good or interesting
	Paper is extremely bad
	Paper is in grey area and I can't make up my own mind
	Teaching
	Teaching tool
54a) What percent of the time do you discuss the manuscript with the editor?	
54b) What triggers discussion with the editor?	
	respondent generated categories
	Major problem with the manuscript
	To flag an exceptional article
	Manuscript previously reviewed for another journal
	Question of suitability of manuscript for Annals
	Statistical issue
	Ethical issue
55a. Do you use a template or checklist?	
56) What percent of the time do you refer back to prior reviews?	
57) What percent of the time do you check citations as part of your review?	
58a) What percent of the time would you pull an abstract or an article?	
58b) What percent of the time do you use clinical reference books?	
58c) What percent of the time do you use references on study design or statistics?	
61a) Have you ever seen the Scoring Elements for Review Quality?	
61b) Was it or do you think it would be useful to see these criteria?	
61c) Would it be useful to know your scores?	
62a) What per cent of the time to you access the editor's decision letter?	
62b) Is the editor's decision letter to the author useful?	
62c) Why is the editor's decision letter to the author useful?	
	respondent generated categories
	Curious about agreement with editor.
	Feedback: Helps me understand what the editor is looking for.
	Provides a synthesis of the critique
63a) What per cent of the time to you access the other reviewers' reviews?	
63b) Are the other reviewers' reviews useful to you?	
63c) Why are the other reviewers' reviews useful?	
	respondent generated categories
	Concurrence: Want to know what (how) other reviewers think
	Deficits: Want to know what I missed
63d) Which do you find more useful: the editor's or reviewer's comments?	
64a) Are your reviews different than they were when you first started?	

64a1) In what way have your reviews changed?	
	respondent generated categories
	REVIEW
	- Length of review
	- reviews are shorter and more focused- reviews are more general and focus more on the big issues
	- reviews are longer, more thorough and more detailed
	- Restrict time spent on articles with fatal flaws.
	- Function of review
	Evaluative Function (triage): Editorial direction - more conscious of giving the editor direction.
	Instructional Function (pedagogy): More conscious of providing the author with help in improving the manuscript.
	- Quality of review
	- is better / more mature / sophisticated / more organized / more confident
	- more critical / skeptical / rigorous
	- less critical / more supportive
	MANUSCRIPT
	- Significance
	- more attention to clinical significance
	- Methodology
	- Pay more attention to methodology
	- Less doctrinaire about methodological issues
	- Statistics
	- pay more attention to stats
	- pay less attention to stats
	- References
	- check references more carefully
	- Understandability
	- less emphasis on language, grammar - no longer bother with copy-editing
	- more emphasis on understandability, conciseness and focus.
	- more emphasis on graphics, charts and tables.
	Ethics
	- now comment on ethics
71) What makes a strong article?	
	respondent generated categories
	Clinically relevant
	Original
	Interesting
	Methodological rigor/sound study design
	Well / clearly written
	Clearly stated hypothesis
	Conclusion is well-founded
	Valid Results
	Limitations explained

	Complete Data
	Discussion sticks to findings
	Prospective RCT
	Study integrated into literature.
	References are complete
	EBM
81) What makes a good review?	
	respondent generated categories
	FUNCTION
	Instruction (pedagogy) : provides the author with ways of improving the manuscript
	Evaluation (triage): Editorial direction
	CONTENT
	Big Picture: Identifies major strengths and weaknesses.
	Detail: Should provide a careful detailed critique of flaws.
	Context: Place the study within the context of present knowledge.
	Validity: Should assess the validity of the findings.
	FORM / VALUES / PROCESS
	Tone: Should be positive and constructive
	Thorough: Evidence of thought, effort and time spent
	Advice should be clear and concise
	Review should be fair / unbiased
82) Which is more important:	
	a. Evaluation (triage): To provide the editor with reasons to publish or not to publish
	b. Instruction (pedagogy): To provide the author with useful suggestions for improvement of manuscript.
	c. Both.
83) Which is more important:	
	a. To identify major weaknesses
	b. To identify major strengths
	c. Both
91) What makes a competent peer reviewer?	
	respondent generated categories
	Knowledge of:
	Domain (topic area)
	Study design and methodological
	Statistical
	Critical review
	Know what will be of interest to the reader
	Knowledge of own limitations
	Experience
	Research experience
	Clinical experience
	Critical Review experience
	Review Skills
	Thorough / compulsive / committed
	Good communication skills

	Prompt
	Consistency
	Writes good reviews
	Personal / Intellectual Attributes
	Fair
	Supportive
	Critical mind
	Open minded
	Training
	In literature review / methodology
101) Do you think the peer review process accomplishes what it should?	
101a) Which is?	
	respondent generated categories
	Evaluation (triage): It
	Identifies good papers.
	Identifies bad papers
	Validates the literature / the specialty.
	Instruction (pedagogy): It improves:
	the paper / research
	skills of the researcher/authors
	Insure Fairness
	Corporate wisdom (three heads are better than one)
	Nil alternatives (democracy analogy)
103) What are the drawbacks of the peer review process?	
	respondent generated categories
	The RESEARCH
	Judgment
	Good science does not get published
	Poor science gets published
	The wrong kind of science gets published
	The REVIEWER
	Selection Biases
	Reviewer variability – too much variation in abilities, skills and results
	Limited reviewer pool – too little variation
	Ethical defects
	Bias
	Politics / conflicts of interest / petty jealousy
	Plagiarism – possibility of ideas being stolen
	Knowledge / Skills defects
	Reviewer lacks knowledge
	Reviewer lacks critical skills
	Reviewer doesn't put necessary effort into it
	The REVIEW
	Communication errors
	Mean Spirited Tone - review is not constructive
	Technical limitations
	Overly subjective and arbitrary
	Judgment

	Hypercritical: overly critical, nitpicking
	Discrepant
	The PROCESS of peer review
	Time
	Turnaround time: Delays in publication
	Time consuming for reviewers
	Standards
	Are too low
	Are too high
104) How would you improve the peer review process?	
	respondent generated categories
	The PROCESS of peer review
	Fine the way it is.
	Time
	Accelerate turnaround times
	Time consuming > Pay reviewers
	Standards
	Standardization: Editors should be clearer about what they want.
	The REVIEWER-
	Selection
	Expand pool of reviewers
	Change selection criteria for reviewers
	Decrease pool of reviewers
	Ethics
	Declarations of interest (esp. funding sources) for both authors and reviewers
	Education
	More Feedback
	Courses
	Checklists or templates
	Mentors
	The REVIEW
	Communication
	Mean-spirited reviews.
	Reviews should be published along with article
105) Do you think reviewing should continue to be 'blinded'?	
105a) Why or why not?	
	respondent generated categories
	- Knowing provenance adversely effects the review.
	- Knowing provenance has little effect on the review.
	- Knowing provenance improves the review.
	- Unblinding would be useful in increasing direct dialogue between reviewers and authors.
111a) Does reviewing take up an inordinate amount of your time?	
112a) Are the rewards of peer reviewing sufficient?	
112b) What do you think you gain by reviewing?	
	respondent generated categories
	Education and Skills

	Increases domain knowledge
	Improves research skills
	Improves critical skills
	Improves manuscript writing skills
	Improves teaching skills
	Improves clinical skills
	Career Advancement
	Advances career
	Status / prestige
	Increases networking possibilities
	Personal satisfaction
	Enjoyment
	Duty
	Helps shape discipline
	Gives insights into the process of science
113a) Should reviewers be paid?	
113b) How much?	
	respondent generated categories
	- Don't know
	- Honorarium - $50 to $100
	- More than token, but less than clinical time - $101 - 200
	- Equal to clinical time - $201 – 300/hr.
113c) What advantages or disadvantages can you see to the payment of reviewers?	
	respondent generated categories
	Advantages
	Improve quality of reviews
	Compensates for time
	Increase number of reviewers
	Gives value to the work.
	Gives the editor some control over quality of reviews and promptness
	Disadvantages
	Financially untenable for journal
	Moral - puts motivation in question
	Quality - Encourages mediocrity
	Conflicts of interest / bias
	Ceases to be truly 'peer' review
114c) How else might reviewers be rewarded?	
	respondent generated categories
	Professional
	Increase weighting in performance evaluations
	Promotion to editorial board
	Appreciation
	Listing in journal
	T shirts, hats, pens etc.
	Certificate/letter from Annals
	Dinner, lunch or reception at conference
	Thank-you e-mail

	Book tokens / gift certificates
	Awards
	Awards for excellence
	Education
	More feedback – to help improve my own skills
	Course on peer review skills
	Special rates for conferences
	Free subscription to a journal
	CME credits
	Research
	Priority for getting own work published in journal.
121) Do you have any other comments about the review process that that you would like to share?	