

**The Hawthorne effect in journal peer review**

**Lutz Bornmann\***

\* Max Planck Society, Administrative Headquarters, Hofgartenstr. 8, 80539 Munich, E-mail:

bornmann@gv.mpg.de, Phone: +49 89 2108 1265

Abstract:

Purpose – This paper aims to look at the Hawthorne effect in editorial peer review.

Design/methodology/approach – Discusses the quality evaluation of refereed scholarly journals.

Findings – A key finding of this research was that in the peer review process of one and the same manuscript, reviewers or editors, respectively, arrive at different judgments. This phenomenon is named as “Hawthorne effect” because the different judgements are dependent on the specific conditions under which the peer review process at the individual journals takes place.

Originality/value – Provides a discussion on the quality evaluation of scholarly journals.

Keywords: Editorial peer review; Manuscript rejection; Hawthorne effect

## Introduction

Peer review is quintessential in science (Benda & Engels, 2011). Since almost all research articles published in journals undergo the peer review process, it is an essential extension of the scientific process (Callaham & McCulloch, 2011). “An article in a reputable journal does not merely represent the opinions of its author; it bears the *imprimatur* of scientific authenticity, as given to it by the editor and the referees he may have consulted. The referee is the lynchpin about which the whole business of Science is pivoted” (Ziman, 1968, p. 148). Peers asked to evaluate manuscripts take on the responsibility for assuring high standards in various research disciplines. “Equals active in the same field are said to be in the best position to know whether quality standards have been met and a contribution to knowledge made” (Eisenhart, 2002, p. 241). Peer evaluation thus entails a process by which a selective jury of equal’s active in a given scientific field convenes to evaluate scientific outcomes.

In journal peer review, reviewers sought by the editor normally provide the editor with a written review and overall publication recommendation. “The editor, on the basis of the reviews and his or her own evaluation, decides to reject the submission, seek further review, ask the author to revise the manuscript in response to suggestions by the reviewers and the editor, or accept the manuscript” (Jayasinghe, Marsh, & Bond, 2001, p. 344). Most of the studies examining the relation of reviewers’ (overall) ratings and editors’ decisions on submissions at single journals have found that the reviewers’ ratings are highly correlated with the editors’ final decisions (Bakanic, McPhail, & Simon, 1987; Bornmann & Daniel, 2008a; Fogg & Fiske, 1993; Lock, 1985; Petty & Fleming, 1999; Sternberg, Hojjat, Brigockas, & Grigorenko, 1997; Zuckerman & Merton, 1971). That means editors render their decisions on manuscripts dependently on the judgments of the reviewers.

As a proponent of peer review Abelson writes (1980): “The most important and effective mechanism for attaining good standards of quality in journals is the peer review system” (p. 62). According to Shatz (2004) journal peer review “motivates scholars to produce their best, provides feedback that substantially improves work which is submitted, and enables scholars to identify products they will find worth reading” (p. 30). For Shatz (2004) “probably the most frequently voiced concern about the peer review system is the potential for bias and lack of objectivity on the part of reviewers” (p. 36). Biases enter peer review, when factors that are independent on the quality of a manuscript and are functionally irrelevant to research correlate statistically with the judgment of reviewers or the decision of the editors (Marsh, Jayasinghe, & Bond, 2008; Weller, 2002), such as:

- Features of authors, e.g. academic status,
- Features of research reported in a manuscript, e.g. the sub-field or the statistical significance of results,
- Features of the reviewers, e.g. gender, and/or
- Features of the journal’s peer review process, i.e. the Hawthorne effect introduced here.

The various overviews of the peer review research literature (see e.g. Bornmann, 2011; de Vries, Marschall, & Stein, 2009; Hojat, Gonnella, & Caellegh, 2003; Martin, 2000; Owen, 1982; Pruthi, Jain, Wahid, Mehra, & Nabi, 1997; Ross, 1980; Sharp, 1990; Wood & Wessely, 2003) have named up to 25 different sources of bias (especially features of authors) that can potentially endanger the fairness of the peer review process. These potential sources of bias can be divided into positive or negative biases: “that is, a bias may lead to a more negative evaluation of an article than the referee would give were it not for the biasing factor, or it may lead to a more positive evaluation, in which case we may speak of a ‘halo effect,’ whereby the quality of a work is exaggerated upward by the

appraiser. Or the bias may make no difference: a biased evaluation might be identical with what an unbiased evaluation would yield” (Shatz, 2004, p. 36).

## The Hawthorne effect – a possible bias in journal peer review

The goal of this paper is to introduce the Hawthorne effect as a phenomenon in peer review which can lead to a biased evaluation of a manuscript: One and the same manuscript is differently assessed by peers and/or editors in dependence on the environment (i.e., the journal). Originally, the Hawthorne effect is a phenomenon of group-based observational studies, which was discovered in experiments undertaken in the Hawthorne factories (Illinois, USA). Participants of a study changed their natural behaviour, if they knew that they take part in a study and are observed. “Careful studies of this wiring group showed marked increases in production which were related only to the special social position and social treatment they received. Thus it was the ‘artificial’ social aspects of the experimental conditions set up for measurement which produced the increases in group productivity” (French, 1953) (cited by Chiesa & Hobbs, 2008, p. 68). The bias is introduced into the experiment by certain environmental aspects of the study itself. In the current literature, such an influence is named as Hawthorne effect.

In the context of journal peer review, the term “Hawthorne effect” was firstly used by van Rooyen, Godlee, Evans, Smith, and Black (1998). They investigated differences in reviewing in dependence on blinded (uninformed about the submitting authors) and un-blinded (informed) reviewers. Their results show that “there was no evidence of any difference between masked un-blinded and uninformed reviewers and therefore no detectable Hawthorne effect” (van Rooyen, et al., 1998, p. 236). Van Rooyen, Godlee, Evans, Smith, and Black (1998) used the term “Hawthorne effect” to describe the influence of the peer review process’ features (i.e., blinded or un-blinded reviewing) on

the evaluation of manuscripts which may influence the quality evaluation. In this paper the Hawthorne effect is formulated more generally by claiming that the evaluation of a submitted manuscript is dependent on the features of a journal's peer review process: Blinded or un-blinded reviewers are one feature of the process which may influence the evaluation.

Pre-publication peer review of manuscripts is generally seen to ensure that only (high) quality research is published (either by rejecting or improving sub-standard manuscripts) (Lipworth, Kerridge, Carter, & Little, 2011). However, it is not only the scientific quality of a manuscript which results in an acceptance or rejection decision, but also journal specific factors. The findings of the studies presented in the following show that one and the same manuscript is differently assessed by different journals. The studies have investigated the fate of manuscripts rejected by one journal and published by another journal. Bornmann and Daniel (2008a, 2008b) could determine for the journal *Angewandte Chemie* that no alterations or only minor alterations were made to approximately three-quarters of the rejected manuscripts for publication elsewhere. If general quality standards were valid at both journals one and the same manuscript would receive the same editorial decision. Since the peer review process of different journals may take a variety of forms (because of, e.g., the use of blinded or un-blinded reviewers, the higher or lower reputation of the journal, or the disciplinary orientation), the handling of the manuscripts is different. The phenomenon of the different handling – independent on the quality of the manuscript – is named here as Hawthorne effect because journal specific factors and not the quality drive the editorial decision.

## The publication of rejected manuscripts elsewhere

In journal peer review “a rejection usually does not kill a paper ... a rejected paper usually finds life at another journal” (Gans & Shepherd, 1994, p. 177). Already in the 1980’s Abelson (1980) reported that almost all of the manuscripts rejected by the journal *Science* were published later in other journals. For manuscripts rejected by the journal *Angewandte Chemie* in the year 1984, Daniel (1993) determined a percentage of 71%; for manuscripts rejected by this journal at the beginning of the year 2000, Bornmann and Daniel (2008a, 2008b) found 95% published later elsewhere. Other studies on the fate of manuscripts rejected by a journal report percentages ranging from 28% to 85% (Weller, 2002). The most recent study on the journal *Radiology* found that 79% of the rejected manuscripts were subsequently published elsewhere (Khosla, McDonald, Bornmann, & Kallmes, 2011).

Some studies on the fate of rejected manuscripts have also gauged the quality of journals that accepted rejected manuscripts. In a literature review covering research on journal peer review, Weller (2002) cites five studies, (Chew, 1991; Cronin & McKenzie, 1992; Gordon, 1984; Weller, 1996; Whitman & Eyre, 1985), which have ranked the quality of the rejecting and the later accepting journals mostly by means of the Journal Impact Factors (JIF, provided by Thomson Reuters in the Journal Citation Reports). The JIF is the average number of times papers from the journal published in the past two years (e.g. 2005 and 2006) have been cited in the JCR year (e.g. 2007). Seven further studies, which are not included in the literature review by Weller (2002), have been published by Bornmann and Daniel (2008a, 2008b), Daniel (1993), Lock (1985), McDonald, Cloft, and Kallmes (2007), Opthof, Furstner, van Geer, and Coronel (2000), Pulverer (2010) and Ray, Berkwits, and Davidoff (2000). In the total of twelve studies, between 0% (Daniel, 1993) and 70% (Gordon, 1984) of the rejected manuscripts in one higher quality journal could be researched. The results of these studies show that rejected manuscripts are frequently published (but not always) in lower impact journals. It seems to be that the higher the quality ranking of the rejecting journal is,

the lower the chance is that a rejected manuscript will appear in another journal ranked as higher quality.

## Discussion

The results of the studies investigating the fate of rejected manuscripts demonstrate that in the peer review process of one and the same manuscript, reviewers or editors, respectively, arrive at different judgments: Many manuscripts that are rejected by a journal (through a peer review process) are then accepted by another journal (through a peer review process). This finding indicates that manuscript review is *not* only based on generally valid quality criteria that a scientific work can fulfil (acceptance) or not fulfil (rejection); the review (or the outcome of the review) seems also to be dependent on journal specific factors under which the peer review process at the individual journals takes place (Bornmann, 2008). It is proposed here that this phenomenon is named as “Hawthorne effect.” The effect is probably boosted in the current science system through the increasing shift of the use of papers as an evaluative measure of scientific performance:

“Because of the intense pressure to get into a handful of top journals, instead of sending less-than-ground-breaking work to second- or third-tier journals, more scientists are first sending their work to elite publications, where they often clearly don’t belong. Consequently, across the board, editors at top-tier journals say they are receiving more submissions every year, leading in many cases to more rejections, appeals, and complaints about the system overall” (McCook, 2006, p. 26).

The Hawthorne effect can be aligned with the Oppenheim effect which was introduced by Gorman (2007). The formulation of this effect is based on the following incident: “Charles Oppenheim got a humbling insight into the reality of academic publishing recently. The professor of information science at Loughborough University, a dominant figure in his field, had submitted an article to a

respected journal. ‘Some time later a very good friend of mine got in touch with me. This friend had been sent my article to referee. He thought that, in all honesty, he could not referee it because we were so close and he telephoned the editor to say that he would have to find someone else. The editor responded: ‘Don’t be silly, this is Charles Oppenheim. We both know we are going to publish it anyway. This is really just a formal exercise, so could you just go through the motions?’ Oppenheim chuckles: ‘The editor had kind of given the game away. Now I know I can send him any old rubbish and get it in’” (Bunting, 2005). The Oppenheim effect describes the phenomenon that the author and not the manuscript itself is the criterion in determining quality to the submission and is the base for the editorial decision.

For Cole (1992) the results of peer review studies suggest that the evaluation of scientific work is influenced by a complex interaction between (1) universalistic factors, such as scientific merit, and (2) social aspects. Cole (1992) assumes that there is no way to objectively evaluate new scientific work. In a similar way, Ziman (2000) finds the influence of non-cognitive factors on academic cycles of credibility and credit quite substantial in the more highly structured social world of post-academic science (see here also Lipworth, et al., 2011). Phenomenon, like the Hawthorne or Oppenheim effect, should not be a surprise therefore.

## References

- Abelson, P. H. (1980). Scientific communication. *Science*, 209(4452), 60-62.
- Anon. (2011). There's a time to be critical. *Nature*, 473(7347), 253.
- Bakanic, V., McPhail, C., & Simon, R. J. (1987). The manuscript review and decision-making process. *American Sociological Review*, 52(5), 631-642.
- Benda, W. G. G., & Engels, T. C. E. (2011). The predictive validity of peer review: a selective review of the judgmental forecasting qualities of peers, and implications for innovation in science. *International Journal of Forecasting*, 27(1), 166-182. doi: 10.1016/j.ijforecast.2010.03.003.
- Bornmann, L. (2008). Scientific peer review. An analysis of the peer review process from the perspective of sociology of science theories. *Human Architecture - Journal of the Sociology of Self-Knowledge*, 6(2), 23-38.
- Bornmann, L. (2011). Scientific peer review. *Annual Review of Information Science and Technology*, 45, 199-245.
- Bornmann, L., & Daniel, H.-D. (2008a). The effectiveness of the peer review process: inter-referee agreement and predictive validity of manuscript refereeing at *Angewandte Chemie*. *Angewandte Chemie International Edition*, 47(38), 7173-7178. doi: 10.1002/anie.200800513.
- Bornmann, L., & Daniel, H.-D. (2008b). Selecting manuscripts for a high impact journal through peer review: a citation analysis of Communications that were accepted by *Angewandte Chemie International Edition*, or rejected but published elsewhere. *Journal of the American Society for Information Science and Technology*, 59(11), 1841-1852. doi: 10.1002/asi.20901.
- Bunting, C. (2005, 25 February). Early careers spent grinding teeth, not cutting them. *Times Higher Education Supplement*, 18.
- Callaham, M., & McCulloch, C. (2011). Longitudinal trends in the performance of scientific peer reviewers. [Proceedings Paper]. *Annals of Emergency Medicine*, 57(2), 141-148. doi: 10.1016/j.annemergmed.2010.07.027.
- Chew, F. S. (1991). Fate of manuscripts rejected for publication in the AJR. *American Journal of Roentgenology*, 156(3), 627-632.
- Chiesa, M., & Hobbs, S. (2008). Making sense of social research: how useful is the Hawthorne Effect? [Article]. *European Journal of Social Psychology*, 38(1), 67-74. doi: 10.1002/ejsp.401.
- Cole, S. (1992). *Making science. Between nature and society*. Cambridge, MA, USA: Harvard University Press.
- Cronin, B., & McKenzie, G. (1992). The trajectory of rejection. *Journal of Documentation*, 48(3), 310-317.
- Daniel, H.-D. (1993). *Guardians of science. Fairness and reliability of peer review*. Weinheim, Germany: Wiley-VCH.
- de Vries, D. R., Marschall, E. A., & Stein, R. A. (2009). Exploring the peer review process: what is it, does it work, and can it be improved? *Fisheries*, 34(6), 270-279.
- Eisenhart, M. (2002). The paradox of peer review: admitting too much or allowing too little? *Research in Science Education*, 32(2), 241-255.
- Fogg, L., & Fiske, D. W. (1993). Foretelling the judgments of reviewers and editors. *American Psychologist*, 48(3), 293-294.
- French, J. R. P. (1953). Experiments in field settings. In L. Festinger & D. Katz (Eds.), *Research methods in the behavioral sciences* (pp. 98-135). New York, NY, USA: Holt, Rinehart and Winston.
- Gans, J. S., & Shepherd, G. B. (1994). How are the mighty fallen - rejected classic articles by leading economists. *Journal of Economic Perspectives*, 8(1), 165-179.

- Gordon, M. D. (1984). How authors select journals - a test of the reward maximization model of submission behavior. *Social Studies of Science*, 14(1), 27-43.
- Gorman, G. E. (2007). The Oppenheim effect in scholarly journal publishing. *Online Information Review*, 31(4), 417-419. doi: Doi 10.1108/14684520710780386.
- Hojat, M., Gonnella, J. S., & Caellegh, A. S. (2003). Impartial judgment by the "gatekeepers" of science: fallibility and accountability in the peer review process. *Advances in Health Sciences Education*, 8(1), 75-96.
- Jayasinghe, U. W., Marsh, H. W., & Bond, N. (2001). Peer review in the funding of research in higher education: the Australian experience. *Educational Evaluation and Policy Analysis*, 23(4), 343-346.
- Khosla, A., McDonald, R. J., Bornmann, L., & Kallmes, D. F. (2011). Getting to yes: the fate of neuroradiology manuscripts rejected by *Radiology* over a 2-year period. *Radiology*, 260(1), 3-5. doi: 10.1148/radiol.11110490.
- Lipworth, W. L., Kerridge, I. H., Carter, S. M., & Little, M. (2011). Journal peer review in context: a qualitative study of the social and subjective dimensions of manuscript review in biomedical publishing. [Review]. *Social Science & Medicine*, 72(7), 1056-1063. doi: 10.1016/j.socscimed.2011.02.002.
- Lock, S. (1985). *A difficult balance: editorial peer review in medicine*. Philadelphia, PA, USA: ISI Press.
- Marsh, H. W., Jayasinghe, U. W., & Bond, N. W. (2008). Improving the peer-review process for grant applications - reliability, validity, bias, and generalizability. *American Psychologist*, 63(3), 160-168. doi: Doi 10.1037/0003-066x.63.3.160.
- Martin, B. (2000). Research grants: problems and options. *Australian Universities' Review*, 43(2), 17-22.
- McCook, A. (2006). Is peer review broken? *The Scientist*, 20(2), 26.
- McDonald, R. J., Cloft, H. J., & Kallmes, D. F. (2007). Fate of submitted manuscripts rejected from the *American Journal of Neuroradiology*: outcomes and commentary. *American Journal of Neuroradiology*, 28(8), 1430-1434. doi: 10.3174/Ajnr.A0766.
- Ophthof, T., Furstner, F., van Geer, M., & Coronel, R. (2000). Regrets or no regrets? No regrets! The fate of rejected manuscripts. *Cardiovascular Research*, 45(1), 255-258.
- Owen, R. (1982). Reader bias. *Journal of the American Medical Association*, 247(18), 2533-2534.
- Petty, R. E., & Fleming, M. A. (1999). The review process at *PSPB*: correlates of interreviewer agreement and manuscript acceptance. *Personality and Social Psychology Bulletin*, 25(2), 188-203.
- Pruthi, S., Jain, A., Wahid, A., Mehra, K., & Nabi, S. A. (1997). Scientific community and peer review system - a case study of a central government funding scheme in India. *Journal of Scientific & Industrial Research*, 56(7), 398-407.
- Pulverer, B. (2010). Transparency showcases strength of peer review. [10.1038/468029a]. *Nature*, 468(7320), 29-31.
- Ray, J., Berkwitz, M., & Davidoff, F. (2000). The fate of manuscripts rejected by a general medical journal. *American Journal of Medicine*, 109(2), 131-135.
- Ross, P. F. (1980). *The sciences' self-management: manuscript refereeing, peer review, and goals in science*. Lincoln, MA, USA: The Ross Company.
- Sharp, D. W. (1990). What can and should be done to reduce publication bias - the perspective of an editor. *Journal of the American Medical Association*, 263(10), 1390-1391.
- Shatz, D. (2004). *Peer review: a critical inquiry*. Lanham, MD, USA: Rowman & Littlefield.
- Sternberg, R. J., Hojjat, M., Brigockas, M. G., & Grigorenko, E. L. (1997). Getting in: criteria for acceptance of manuscripts in *Psychological Bulletin*, 1993-1996. *Psychological Bulletin*, 121(2), 321-323.

- van Rooyen, S., Godlee, F., Evans, S., Smith, R., & Black, N. (1998). Effect of blinding and unmasking on the quality of peer review - A randomized trial. *Journal of the American Medical Association*, 280(3), 234-237.
- Weller, A. C. (1996). Editorial peer review: a comparison of authors publishing in two groups of US medical journals. *Bulletin of the Medical Library Association*, 84(3), 359-366.
- Weller, A. C. (2002). *Editorial peer review: its strengths and weaknesses*. Medford, NJ, USA: Information Today, Inc.
- Whitman, N., & Eyre, S. (1985). The pattern of publishing previously rejected articles in selected journals. *Family Medicine*, 17(1), 26-28.
- Wood, F. Q., & Wessely, S. (2003). Peer review of grant applications: a systematic review. In F. Godlee & T. Jefferson (Eds.), *Peer review in health sciences* (2nd ed., pp. 14-44). London, UK: BMJ Books.
- Ziman, J. (1968). *Public knowledge: an essay concerning the social dimension of science*. Cambridge, UK: Cambridge University Press.
- Ziman, J. (2000). *Real science. What it is, and what it means*. Cambridge, UK: Cambridge University Press.
- Zuckerman, H., & Merton, R. K. (1971). Patterns of evaluation in science: institutionalisation, structure and functions of the referee system. *Minerva*, 9(1), 66-100-100. doi: 10.1007/bf01553188.