What can new technology tell us about the reviewing process for journal submissions in BJET?

James Hartley, Guillaume Cabanac

What can new technology tell us about the reviewing process for journal submissions in *BJET*?

James Hartley & Guillaume Cabanac

James Hartley is Emeritus Professor of Psychology at Keele University, UK. He is widely known for his work on student learning, text design, and academic writing. Address for correspondence: James Hartley, School of Psychology, Keele University, Staffordshire, ST5 5BG, UK.
Email: j.hartley@keele.ac.uk.

Guillaume Cabanac is Associate Professor of Computer Science at the University of Toulouse, France. His academic interests lie in Information Science with a focus on Information Retrieval and Scientometrics.
Email: guillaume.cabanac@univ-tlse3.fr

Abstract

New technology allows us to collect rich and voluminous data about refereeing — something that was not possible before. In this quantitative study we consider data about the dates of sending (i) suitable articles to *BJET*, (ii) the articles to referees, and (iii) when the referees' reports are returned to the editor.

To do this we examined data from over 11,000 reports collected between 2005 (when *BJET* began using ScholarOne as its online editorial manager) and October 2014. We found most of the submissions and responses were made in the earlier part of the week, with some evidence of academics (and the editor) working on Sundays. We also found modest agreement between referees on the quality of individual papers.

Practitioner's notes

What is already known about this topic?

- Editors vary in their methods of locating referees, and in their instructions to them concerning the amount of time they have for review, and how they should complete such reviews, e.g., by using comments, rating scales or check boxes. However, little is known about the effectiveness of these different methods.
- Reviewers vary in their time to referee papers, but again little is known about the causes of such variation and how it might affect their reports.
- There is only modest agreement between reviewers when they referee individual papers for inclusion in a journal.
- And there is little research on the effectiveness of different practices used and advocated by editors.

What this paper adds

- A review of the current literature in this field.
- A picture of editorial and reviewer practice at *BJET.* The data suggest that the editor of *BJET* assigns more manuscripts for review at the beginning than at the end of the week (some 20%), and that, similarly, reviewers and authors also submit their manuscripts early in the week. There is a blip on Sundays, however, showing that the editor, and 10% of the authors and referees work on Sundays. These different practices, however, appear to have little effect on the referees' decisions for *BJET*. About 10% of the submitted papers are accepted for publication.

Implications for practice and policy

- Such studies need to be repeated to find out whether or not such findings occur with different-sized journals, different subject matters, and with editors using more traditional methods of selecting referees than *BJET*.
- Research with data from electronic publishing systems will help us gain information on many issues to do with academic publishing, such as how far referees agree, the critical number of referees required, and the characteristics of effective reviewers.

Introduction

New technology allows journal editors and publishers to collect and assess more data about submissions, referees' reports, and referees' recommendations than ever before (Fox & Burns, 2015; Hartley, 2012; Hartley & Cowan (submitted)). More can now be done, for example, on the following questions (just to list a few):

- Do journals using different refereeing systems report similar findings (Bormann & Daniel, 2010)?
- Do journals using only two referees (as opposed to three) report similar findings (Schultz, 2010)?
- Do journals offering incentives to reviewers to submit their reviews more quickly speed up reviewing times (Chetty, Saez & Sándor, 2014)?
- Do the private comments of referees to the editors influence the editor's judgements (Anonymous, 2010; Bornmann, Weymuth & Daniel, 2010)?
- Are papers reviewed 'blind' (i.e. with the authors anonymised) treated more harshly by reviewers and, conversely, are papers subject to 'open review' treated more leniently? (Hames, 2007)
- Are papers reviewed by seasoned academics treated less harshly than those reviewed by novices (Callaham & Tercier, 2007)?
- Are papers with multiple authors cited more than papers with single authors (Gazni & Thelwall, 2014; Hartley & Cabanac, 2015)?

- Are male referees (and editors) are more critical of female authors and vice versa (Ceci, Ginther, Kahn & Williams, 2014; Chibnik 2014)?
- Are papers submitted to journals citing the journal's editor in their text treated more kindly by the editors than papers that do not do so (Frandsen & Nicolaisen, 2011; García, Rodriguez-Sanchez & Fdez-Valdivia, 2015; Sugimoto & Cronin, 2013)?

Using electronic databases we can now find out much more about how journals operate. In this present study, for example, we were able to examine a dataset proving information on 11,437 reports to the *British Journal of Educational Technology* (*BJET*) and on 4,024 manuscripts, courtesy of the editor and Wiley press.

For readers unaware of the peer review process of *BJET*, we need to note that *BJET* uses "peer choice" as its means of obtaining referees (Rushby, 2009). In this case the editor regularly distributes electronic copies of the abstracts from some 30 submissions or so to a panel of volunteer reviewers. The names of the authors are not revealed at this stage. This panel has over 500 members, but not all of them are active. Individual members of the panel can bid and prioritize up to three papers from this list that they would like to review. (If more than three referees bid to review the same paper, the editor decides which ones to ask.) These volunteers are then given up to 30 days to review their selected paper(s) (which now have named authors), and a reminder is sent after some 20 days or more if a review has not been received. The referees' reports are anonymous as far as the authors are concerned. If no review is forthcoming, then the paper is re-allocated to another referee, or placed back in the next list of submissions, with an indication that this particular paper urgently needs reviewers. Peer-choice has an advantage for the editor in that he does not need to select different individual reviewers for every paper received, and an advantage for reviewers in that they choose to evaluate the papers on topics they feel most comfortable with.

Data

The dataset generated by ScholarOne comprises 11,457 lines containing the information shown in Figure 1. There are four main fields:

1. The manuscript ID (e.g., BJET-0042-Mar-2005-OMS) showing the month and year of the submission (March 2005) and the rank number of the manuscript in the year (42th paper received in 2005). Revisions are marked by suffixing the manuscript ID with R1, R2, and so on (e.g., BJET-0042-Mar-2005-OMS.R1 and BJET-0042-July-2005-OMS.R2). The type of the manuscript is one of the following:
    a. OMS: Original Manuscript (88.6%)
    b. COL: Colloquium (9.4%) (Non peer-reviewed papers. Not now available)
    c. REV: Review (1.6%)
    d. LET: Letter to the editor (0.3%)
    e. OBY: Obituary (0.1%)
2. Timestamps in GMT timezone of various events.

a. When the paper was submitted by the authors (Submission Date).

b. When the paper was sent for review to the referees by the editor, one line per referee (Date Reviewer Assigned).

c. When the review was returned by the referee (Date Score Sheet Completed).

Note that the dataset does not feature the location of ScholarOne users. Papers submitted by Asian authors (GMT+10) might well appear to have an earlier date with a GMT timestamp, while those from North America (GMT-8) might have a later date.

3. The reviewer's recommendation:

    a. Accept
    b. Minor revision
    c. Major revision
    d. Reject and resubmit
    e. Reject

4. The status of the paper:

- Accept
- Assign Reviewers
- Awaiting Reviewer Scores
- Complete EO Processing
- Major revision
- Make EIC Decision
- Minor Revision
- Reject
- Reject and Re-submit
- Revision
- Select Reviewers
- Withdrawn

Full reviewer report

Manuscript ID	Manuscript Status	Submission Date	Date Reviewer Assigned	Date Score Sheet Completed	Reviewer Recommendation
BJET-0324-Aug-2014-OMS	Make EIC Decision	28 Aug 14 11:02:53 AM	23 Sep 14	21 Oct 14	Reject & Resubmit
BJET-0324-Dec-2008-OMS	Reject	03 Dec 08 5:55:14 PM	15 Dec 08	06 Jan 09	Minor Revision
BJET-0324-Dec-2008-OMS	Reject	03 Dec 08 5:55:14 PM	15 Dec 08	04 Jan 09	Reject
BJET-0324-Jul-2012-COL	Reject	28 Jul 12 12:12:05 AM			
BJET-0324-Nov-2009-OMS	Reject	07 Nov 09 5:59:48 PM	12 Nov 09	04 Dec 09	Reject
BJET-0324-Nov-2009-OMS	Reject	07 Nov 09 5:59:48 PM	23 Dec 09	23 Dec 09	Reject
BJET-0324-Sep-2010-OMS	Reject	24 Sep 10 2:08:43 AM			
BJET-0325-Aug-2011-OMS	Reject	05 Aug 11 4:36:54 AM			
BJET-0325-Aug-2013-OMS	Withdrawn	11 Aug 13 3:41:15 PM			
BJET-0325-Aug-2014-OMS	Major Revision	29 Aug 14 9:08:05 AM			
BJET-0325-Aug-2014-OMS.R1	Awaiting Reviewer Scores	02 Sep 14 4:43:56 AM	07 Oct 14	22 Oct 14	Reject & Resubmit
BJET-0325-Aug-2014-OMS.R1	Awaiting Reviewer Scores	02 Sep 14 4:43:56 AM	01 Oct 14		
BJET-0325-Aug-2014-OMS.R1	Awaiting Reviewer Scores	02 Sep 14 4:43:56 AM	24 Oct 14		
BJET-0325-Aug-2014-OMS.R1	Awaiting Reviewer Scores	02 Sep 14 4:43:56 AM			
BJET-0325-Dec-2008-COL	Accept	03 Dec 08 8:50:06 PM			
BJET-0325-Jul-2012-OMS	Major Revision	28 Jul 12 3:21:18 AM			
BJET-0325-Nov-2009-OMS	Reject & Resubmit	08 Nov 09 5:58:20 AM	26 Dec 09	09 Jan 10	Major Revision

Figure 1 - Sample of the dataset produced by ScholarOne Manuscripts

In this paper we report on our analyses of these data in terms of the days of the week in which articles are received, the days of the week in which reviewers are assigned to them, the time taken (in days) to receive the reviews, and the reviewers' recommendations (i.e., 1, 2, and 3 above). The original *BJET* dataset and our calculations are published along with this paper as Online Supporting Information **[electronic link to be inserted by the publisher].**

Results

The dynamics of the BJET peer review process

First we asked, does the study of event dates (i.e., submission, assignment, and review) inform us about the dynamics of peer review in *BJET*? Is the workload uniformly distributed Monday to Friday, or are there more subtle variations?

Figure 2 shows the days of the week that authors submitted their papers to *BJET* (either the initial or a revised version, and the type of manuscript). It can be seen that the number of submissions (*N* = 6,130) declines slightly through the week, with Monday being the most frequent (18%). There is a lull on Saturdays (7%), and a slight resurgence on Sundays (9%).

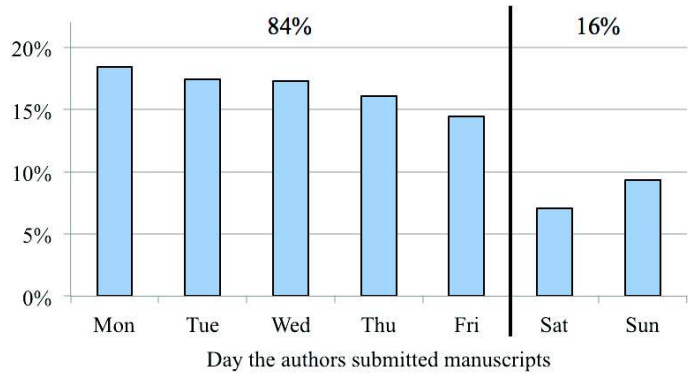

Figure 2 – Distribution of the days of manuscript submission (*N* = 6,130)

Figure 3 shows the days of the week that the editor distributed the submitted papers to reviewers (*N* = 8,373 assignments to one reviewer). It can be seen that 40% were assigned by the beginning of

the week (Mondays and Tuesdays), with reducing numbers towards the end. The day least used for distributing the papers was Saturday (8%), although there was a slight resurgence on Sunday (10%). These figures do not reflect "desk rejects" (i.e., status "Reject" for manuscripts ID that do not contain the review round marker ".R") called by the editor for 33.6% of all submitted papers, as no timestamp is associated with these records.

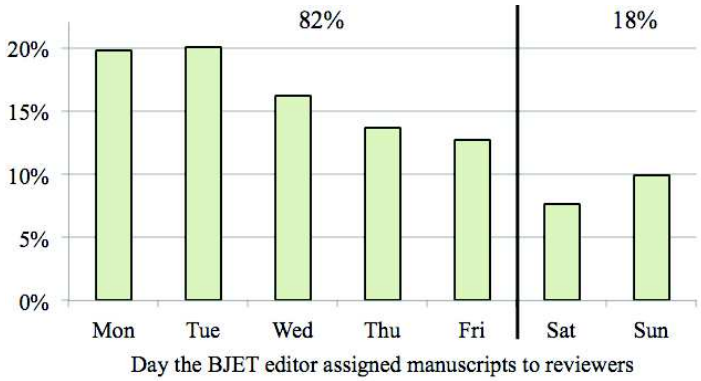

Figure 3 – Distribution of the days of manuscript assignment to reviewers (*N* = 8,373)

Figure 4 shows the days of the week that the referees submitted their reports (*N* = 8,270). This figure shows a similar distribution across weekdays to that of the editor. Half of these reports were submitted early in the first part of the week (Monday to Wednesday), with one-fifth of the reviews posted during the weekends.

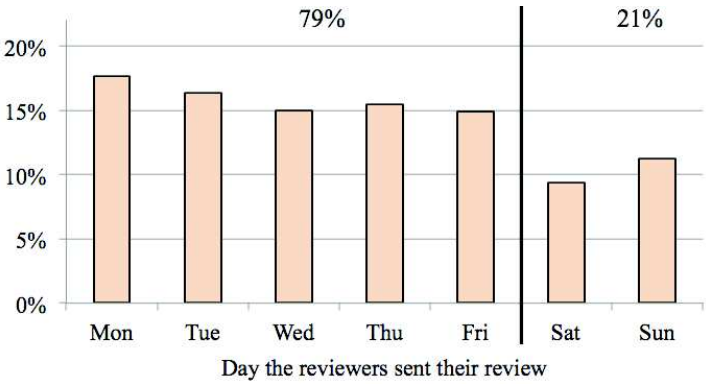

Figure 4 – Distribution of the days when reviews were posted on the editorial manager (*N* = 8,270)

Studying the data from a more long term perspective, Figure 5 shows the pattern of submissions from January 2005–October 2014 (*N* = 6,130 papers of any type, initial submissions and revisions). The yearly number of submissions generally increased with two peaks of growth in 2006 (+70% with respect to 2005) and 2011 (+33% with respect to 2010). The data show no evidence of seasonal submission patterns as claimed by others (e.g., Shalvi, Bass, Handgraaf & De Dreu, 2010). The dip in 2012 probably reflects that Colloquium papers were no longer sought or published, because of their low citation rate and its adverse effect on the impact factor of the journal (personal communication from the editor).

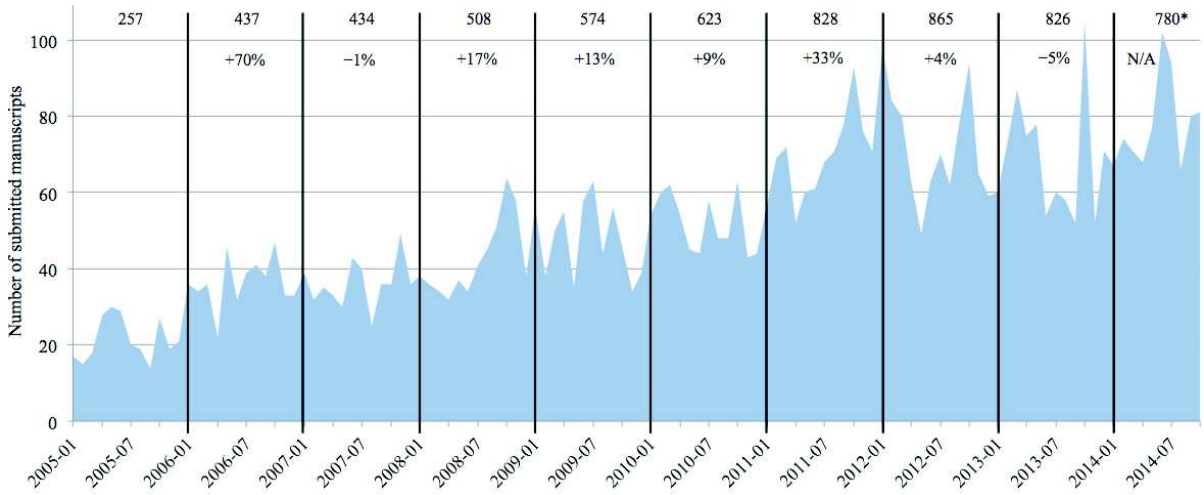

Figure 5 – Distribution of submission dates of manuscripts between January 2005 and October 2014 (*N* = 6,130). Yearly number of submissions and growth with respect to the previous year are shown on top of the graph.

Inter-reviewer agreement

Next we assessed the agreement between reviewers by considering two groups of recommendations: positive ones (i.e., "Accept" and "Minor revision") and negative ones (i.e., "Major revision," "Reject and resubmit," and "Reject"). Here we focused on "OMS" (original manuscripts) initial submissions and revised manuscripts. We computed the κ (kappa) inter-rater agreement coefficient (Fleiss, 1971) between the 2+ reviewers for each such paper (*N* = 3,215). Kappa ranges are usually interpreted according to the nomenclature introduced by Landis & Koch (1977, p. 165): –1.0 ≤ Poor < 0.0 ≤ Slight ≤ 0.2 < Fair ≤ 0.4 < Moderate ≤ 0.6 < Substantial ≤ 0.8 < Perfect ≤ 1.0. Thus there is a "fair agreement" (κ = 0.262) between the reviewers of OMS type papers submitted to the journal (Table 1).

Table 1 – Agreement between the reviewers of original manuscripts (i.e., submission type OMS) for each review turn. The kappa (κ) inter-rater agreement coefficient (Fleiss, 1971) is computed for manuscripts with at least two reviewers assigned.

Review Turn	N manuscripts	N reviewers/manuscript	Inter-reviewer agreement	
			κ value	Interpretation
–	1,778	2.4	0.154	Slight
R1	979	2.5	0.275	Fair
R2	323	2.4	0.191	Slight
R3	105	2.5	0.167	Slight
R4	22	2.6	0.186	Slight
R5	6	2.6	0.074	Slight
R6	2	1.5	– 0.333	Poor
All	3,215	2.4	0.262	Fair

Relationship between length of time for reviews to be completed and the severity of the recommendations

Finally, we looked for a relationship (if any) between the reviewers' recommendations and the number of days for these recommendations to be reported to the Editor. To do this we considered the reviews of the 1,805 OMS manuscripts sent for the first time (i.e., not .R1, .R2, and so on). The boxplots in

Figure 6 show no practical differences in the number of days taken to make a positive vs. a negative recommendation with a median of 23.7 days for both outcomes. Thus we found no relationship between the time taken for the reviews to be completed and the severity of the recommendations.

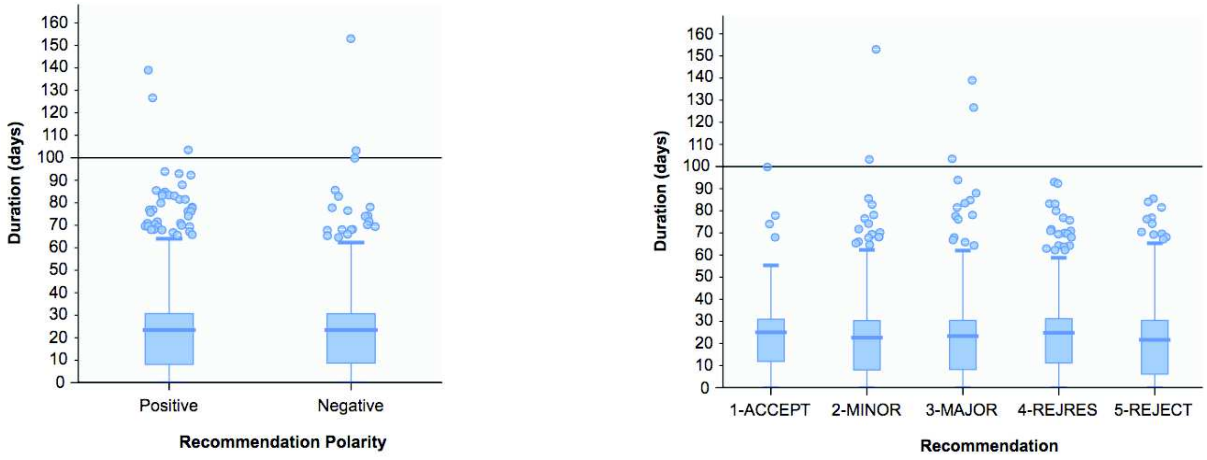

Figure 6 – Number of days to report a positive recommendation (1- ACCEPT or 2 - MINOR REVISION) or a negative one (3 MAJOR REVISION, 4-REJECT AND RESUBMIT, or 5-REJECT) for the 1,805 OMS manuscripts submitted for the first time.

Discussion and Conclusions

The data provided above show no regular patterns in terms of submission, distribution, referee's reports, and recommendations for *BJET*, and there is no evidence to show that referees' reports submitted later are any different from those submitted earlier in terms of their recommendations. (Similar results were reported by Bornmann & Daniel, 2010, p. 73, Table 2).

Nonetheless, our data do provide interesting evidence on other matters. For example, the data shown in Figure 5 provide further evidence *against* the view that it is better to write (in the summer) when it is hot and submit (in the winter) when it is not (Shalvi, Bass, Handgraaf & De Dreu, 2010). Data obtained from other journals also suggest that regular monthly submissions are more typical (Hartley, 2011).

Our results also provide support for another feature of academic writing that suggests that many authors (and editors) work at weekends. Most submissions were made on Monday, and there was strong evidence for authors and editors working on Sundays. These results are in line with those reported by Cabanac and Hartley (2013) and Campos-Arceiz, Koh, and Primack (2013), but they are possibly different in different cultures (Magnone, 2013; Wang, et al., 2012).

New technology thus allows us to explore in far more detail than before the quantitative aspects of journal publishing. Nonetheless, in future, these details need to be illuminated with other forms of analysis – involving qualitative measures. In other words we need to ask referees about why they adopt the procedures that they do, and what they feel about the whole process of reviewing.

Acknowledgements: We are grateful to Wiley Blackwell and Nick Rushby (the editor of *BJET*) for allowing us to use their data, and to Nick Rushby for correcting any mistakes we made in earlier drafts about the procedures used in *BJET*. We are also indebted to our anonymous reviewers for encouraging us to consider different points of view.

References

Anonymous (2010). Referees' quotes – 2010. Environmental Microbiology, *12*, 12, 3303-3304. doi:10.1111/j.1462-2920.2010.02394.x

Bornmann, L. & Daniel, H-D. (2010). How long is the peer review process for journal manuscripts? A case study on *Angewandte Chimie International Edition. Chimia, 64,* 1/2*,* 72-77. doi:10.2533/chimia.2010.72

Bornmann, L., Weymuth, C. & Daniel, H-D. (2010). A content analysis of referees' comments: How do comments on manuscripts rejected by a high-impact journal and later published in a low-impact journal differ? *Scientometrics, 83,* 2, 493-506. doi:10.1007/s11192-009-0011-4

Cabanac, G. & Hartley, J. (2013). Issues of work-life balance among *JASIST* authors and editors. *Journal of the American Society for Information Science and Technology, 64,* 10, 2182-2186. doi:10.1002/asi.22888

Callaham, M.L. & Tercier J. (2007). The relationship of previous training and experience of journal peer reviewers to subsequent review quality. *PLoS Medicine*, *4*, 1, e40. doi:10.1371/journal.pmed.0040040

Campos-Arceiz, A., Koh, L. P. & Primack, R. B. (2013). Are conservation biologists working too hard? [Editorial]. *Biological Conservation*, *166*, 186-190. doi:10.1016/j.biocon.2013.06.029

Ceci, S. J., Ginther, D. K., Kahn, S. & Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science, 15,* 3, 75-141. doi:10.1177/1529100614541236

Chetty, R., Saez, E. & Sándor, L. (2014). What policies increase prosocial behavior? An experiment with referees at the *Journal of Public Economics. Journal of Economic Perspectives, 28*, 3, 169-188. doi:10.1257/jep.28.3.169

Chibnik, M. (2014). From the editor: Gender and citations in *American Anthropologist* [Editorial]. *American Anthropologist, 116,* 3, 493-496. doi:10.1111/aman.12117

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76, *5*, 378-382. doi:10.1037/h0031619

Fox, C. W. & Burns, C. S. (2015). The relationship between manuscript title structure and success: Editorial decisions and citation performance for an ecological journal. *Ecology & Evolution, 5,* 10, 1970-1980. doi:10.1002/ece3.1480

Frandsen, T. F., & Nicolaisen, J. (2011). Praise the bridge that carries you over: Testing the flattery citation hypothesis. *Journal of the American Society for Information Science and Technology, 62,* 5, 807-818. doi:10.1002/asi.21503

García, J., Rodriguez-Sanchez, R. & Fdez-Valdivia, J. (2015). The author-editor game. *Scientometrics, 104,* 1, 361-380. doi:10.1007/s11192-015-1566-x

Gazni, A. & Thelwall, M. (2014). The long term influence of collaboration on citation patterns. *Research Evaluation, 23*, 3, 261-271. doi:10.1093/reseval/rvu014

Hames, I. (2007) *Peer review and manuscript management in scientific journals.* Oxford: Blackwell.

Hartley, J. (2011). Write when you can, and submit when you are ready! *Learned Publishing, 24,* 1, 29-31. doi:10.1087/20110105

Hartley, J. (2012). Refereeing articles in the information age. *British Journal of Educational Technology*, *43*, 3, 520-528. doi:10.1111/j.1467-8535.2011.01211.x

Hartley, J. & Cabanac, G. (2015). An academic odyssey: Writing over time. *Scientometrics*, *103,* 1073-1082. doi:10.1007/s11192-015-1562-1

Hartley, J. & Cowan, J. (submitted). Peer choice – new ways of refereeing and new ways of tracking referees.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 1, 159-174. doi:10.2307/2529310

Magnone, E. (2013). A scientometric look at calendar events. *Journal of Informetrics, 7,* 1, 101-108. doi:10.1016/j.joi.2012.09.006

Rushby, N. (2009). The *BJET* Reviewer panel [Editorial]. *British Journal of Educational Technology*, *40*, 6, 975-979. doi:10.1111/j.1467-8535.2009.01028.x

Schultz, D. M. (2010). Are three heads better than two? How the number of reviewers and editor behaviour affect the rejection rate. *Scientometrics, 84,* 2, 277-292. doi:10.1007/s11192-009-0084-0

Shalvi, S., Baas, M., Handgraaf, J. J. & De Dreu, C. K. W. (2010). Write when hot – submit when not: Seasonal bias in peer review or acceptance? *Learned Publishing, 23,* 2, 117-112. doi:10.1087/20100206

Sugimoto, C. R. & Cronin, B. (2013). Citation gamesmanship: testing for ego bias in peer review. *Scientometrics, 95,* 3, 851-862. doi:10.1007/s11192-012-0845-z

Wang, X., Xu, S., Peng, L., Wang, Z., Wang, C., Zhang, C., & Wang, X. (2012). Exploring scientists' working timetable: Do scientists often work overtime? *Journal of Informetrics, 6*, 4, 655-660. doi:10.1016/j.joi.2012.07.003