Title
Are American astrophysics papers accepted more quickly than others? Part II: correlations with citation rates, subdisciplines, and author numbers

Journal
Scientometrics, 95(1)

ISSN
0138-9130

Authors
Trimble, Virginia
Ceja, Jose A

Publication Date
2013-04-01

DOI
10.1007/s11192-012-0793-7

Scientometrics (2013) 95:45–54
DOI 10.1007/s11192-012-0793-7

Are American astrophysics papers accepted more quickly than others? Part II: correlations with citation rates, subdisciplines, and author numbers

Virginia Trimble · Jose A. Ceja

Received: 20 April 2012 / Published online: 16 June 2012

Abstract We continue the investigation for more than 2,150 astrophysics papers published from July 2007 to June 2008 of various possible correlations among time from submission to acceptance; nationalities of lead authors; numbers of citations to the papers in three years after publication; subdisciplines; and numbers of authors. Paper I found that submissions from American authors were accepted faster than others but by only about 3.8 days out of a median of 105 days. Here we report the following additional relationships: (1) the correlation of citation rate with lag time is weak, the most cited papers having intermediate lag times, (2) citation rates are highest for papers with European and American authors and much smaller for papers from less-developed (etc.) countries, with other prosperous countries in between, (3) citation rates are much larger for currently hot topics (exoplanets, cosmology), than for less hot ones (binary stars, for instance), (4) papers with many authors (seven to more than 100) are more often cited than 1–2 author ones, but this is not linear, and author numbers are not much correlated with lag times, and (5) the lag time for hot topics is about the same as that for less hot topics, which surprised us. Of specific subfields, solar papers are, on average, accepted fastest, quite often within less than 2 months. We don't know why.

Keywords Astronomical journals · Publications · Citations

V. Trimble
Department of Physics and Astronomy,
University of California, Irvine, CA 92697-4575, USA
e-mail: vtrimble@astro.umd.edu

V. Trimble
Las Cumbres Observatory, Santa Barbara, CA 93117, USA

J. A. Ceja (✉)
San Fernando, CA 91340, USA
e-mail: jarmando76@gmail.com

Introduction

Scientists care about a very wide range of aspects of their research, but two fairly important ones (affecting things like job advancement and funding) are how long it takes to get their results published and how often their results are cited by others. Here we address, among other topics, whether these two quantities—lag times from submission to acceptance and citation rates—are correlated.

An earlier paper (Trimble and Ceja 2011, "paper I") looked at about 2,160 astrophysics papers published between July 2007 and June 2008 in the international research journal with the highest citation rates (*Astrophysical Journal*, main journal only) and found that papers submitted by American authors were accepted fastest, but by less than 4 days out of a median of 105 days, surely enough not enough to affect who gets the Nobel Prize, and probably not even enough for anyone to have noticed it. This contrasts with the situation in stem cell research (Aldhaus 2010a, b), where American papers are accepted faster by 24 days out of 83.

An obvious follow-up question is whether outstandingly good papers are accepted faster than average ones, or even below-average ones. We will use numbers of citations to a paper in 3 years after publication as a proxy for the good/average/below-average distinction. The three-year window largely removes the burst of citations that occasionally follows the publication of something manifestly wrong or anyhow out of fashion. It also allows for some smoothing among topics where everybody seems to be in a hurry versus more relaxed ones. And as long as author Ceja had done the work involved in collecting all the citation numbers, author Trimble thought we might as well go ahead and look at some of the other obvious potential correlations.

We will discuss these in the order they are mentioned in the abstract, beginning with average numbers. Figure 1 shows two things. The bar graph and left hand scale present numbers of papers that receive 0, 1, 2, up to more than 100 citations in the 3 years after publication. First quartile, median, and 3rd quartile are 5, 10, and 20 citations per paper. The rising curve and right hand scale display the fraction of all citations going to papers with less than n per paper. Again, half of all citations go to papers receiving 10 or fewer in 3 years (the median), while the mean is larger (15.8) and the mode smaller (4). Only a very small number of papers (16) received more than 100 citations, with the maximum number of citations being 294.

Lag times, citation rates, and countries of origin

Table 1 has what we learned about these, and is rather heavy going. The top line is intervals of time between submission and acceptance in intervals of 0.1 in log time converted to months. The next four lines separate the papers by location of home institutions of the first authors of the papers. Numbers are citations/papers, so the upper left number is 1,080 citations to 66 papers. The fifth line sums papers from all over the world in each time bin, and the bottom lines sum rapid (less than 2 months), average (more than 2 but less than 6.31 months) and slow (more than 6.31 months) papers and also give citations per paper (C/P) for each of these bins. Our first conclusion is, therefore, that both very short and very long interval papers are less cited than the middle range, confirming one of Abt's earlier results (2010). But the difference is not very large in our sample as a whole, 13.9 citations per paper on the wings, versus 17 in the middle. This kind of symmetry is present when the data is divided by regions (as discussed below) except in the case of "third world" countries where citations per paper decrease as lag time increases (we provide a summary of Table 4 in the Appendix).

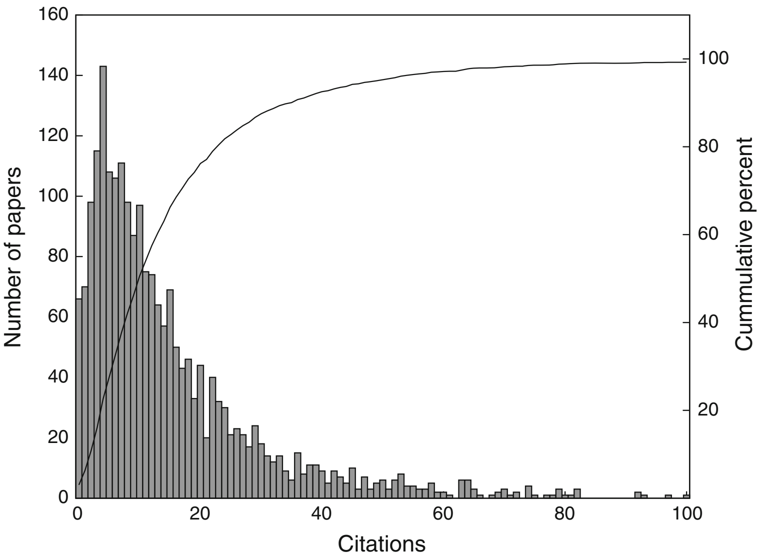

Fig. 1 Distribution of papers versus number of citations for the full data set. The *bar graph* and *left hand axis* show numbers of papers receiving from 0 to 100 or more citations in 3 years. The *curve* and *right hand axis* show fraction of all papers receiving *n* or fewer citations. Both indicate that the median number is about 10 citations per paper, the average about 15, and the mode 4

The right hand column of Table 1 sums the papers by countries of submission. US means United States. Europe means the countries that supported the journal *Astronomy and Astrophysics* and/or the European Southern Observatory at the beginning of 2007. Since then, Argentina, Brazil, and Chile have joined one or both. "Other prosperous" includes Japan, Israel, Canada, Australia, New Zealand, and South Africa, but not South Korea which does not have the long astronomical tradition found in former British Dominions. Finally, developing, less developed, or third world countries include India and China (the largest contributors) but also Latin America, the rest of Africa, South Korea, and a very few papers from Muslim and/or Arabic-speaking countries.

The difference here is conspicuous. Europe and the US lead the citation race, other prosperous countries come in the middle, and papers submitted from developing (etc.) countries pile up the fewest citations. This remains true if you take other slices by subject matter, delay time, or numbers of authors.

Citation rates for subfields of astronomy

It is obvious from Table 2 (abstract item 3) that some parts of astronomy yield papers that are much more frequently cited than others. This is not just a matter of community sizes; there are many more stellar papers (and stellar astronomers) than exoplanet ones. Or at least there were in 2007–2008—in the last few years every institution seems to have developed an exoplanet group. The subtopics were defined as follows:

- Cosmology includes very large scale structure and streaming (observations, formation, and evolution); the correlation of central black hole masses with central velocity dispersions; mergers etc. at redshifts more than about two; data that bear directly on

Table 1 "Grand canonical ensemble" of time from submission to acceptance, citation rates, and countries of origin for all papers for which information was available

	$\Delta t < 1.26$	1.26–1.58	1.58–2.00	2.00–2.51	2.51–3.16	3.16–3.98	3.98–5.01	5.01–6.31	6.31–7.94	7.94–10.00	$\Delta t > 10.00$	Total
US	1,080/66	1,096/74	1,828/117	2,702/147	3,481/188	3,439/196	2,994/160	1,938/112	1,378/89	903/68	1,538/84	22,377/1,301 = 17.2
Europe	323/23	261/23	407/33	815/32	998/59	1,142/62	954/44	846/43	264/14	232/16	414/31	6,656/380 = 17.5
Other prosperous	132/13	65/7	182/16	370/31	668/37	544/39	504/40	379/24	128/11	94/10	239/21	3,305/249 = 13.3
Third world	73/11	151/13	136/18	222/18	135/29	343/35	298/34	109/18	64/17	77/12	79/17	1,687/222 = 7.6
Total	1,608/113	1,573/117	2,553/184	4,109/228	5,282/313	5,468/332	4,750/278	3,272/197	1,834/131	1,306/106	2,270/153	
C/P	14.2	13.4	13.9	18.0	16.9	16.5	17.1	16.6	14.0	12.3	14.8	
Subtotals		5,734/414				22,881/1348				5,410/390		
C/P		13.85				16.97				13.87		
Grand total						34,025/2152						
C/P						15.81						

Vertical columns are by delay time in months (with groups of short, medium, and long at the end)

Lines indicate papers submitted from the US, Europe, other prosperous countries, and less developed countries, as defined in text

Table 2 Numbers of papers, citations, and citations per paper (in 3 years after publication) categorized by subject matter

Topic	Citations	Papers	C/P
Cosmology	9,593	375	25.7
High Energy Astrophysics	7,250	550	13.2
Galaxies	3,366	207	16.3
Star formation/YSOs	2,840	184	15.4
Milky way, ISM, star clusters	2,333	169	13.8
Stars, binaries	2,724	246	11.1
Exoplanets	3,002	140	21.4
Sun, solar system	2,484	244	10.2
Other	445	47	9.5
Total	34,037	2,162	15.7

Clearly cosmology and exoplanets are high impact areas, stars and sun and solar system much less so

how galaxies form and evolve (even if measured at modest z); QSOs and other active galaxies used as cosmological probes; and, of course, measurements of the standard parameters, H_0, Ω_0, and all.

- High Energy Astrophysics includes black holes, neutron stars, cataclysmic variables as X-ray sources, gamma ray bursts (no longer more cited than other parts of the territory cf. Trimble and Ceja 2008), supernovae and their remnants, X-ray binaries, novae (a very few papers), and Active Galaxies considered for their own sake.
- Galaxies means individual ones, but also clusters studied per se, rather than as part of large scale structure.
- Star formation and young stellar objects (YSOs)—the only ambiguity here is the precise line between cloud cores and other kinds of structure in the interstellar medium.
- Milky way includes our own star clusters, stellar populations in clusters in other nearby galaxies, interstellar medium phases and structure not obviously connected with incipient star formation, and associated laboratory chemistry.
- Stars and binaries studied individually, evolutionary tracks and scenarios, chromospheres and coronae (even if X-ray data), pulsations and asteroseismology.
- Exoplanets means discoveries, properties, formation mechanisms, studies of protoplanetary disks (still a hot topic, though not quite so much so as 5 years ago).
- Sun and solar system includes planets, moons, comets, asteroids, interplanetary medium orbiting and around our sun, and pluto.
- "Other" has a few catalogues of objects selected by wavelength etc. and so of several types and a number of theories (e.g., disk instabilities) that could apply to several different types of sources, or to none. This was a small, highly cited category in earlier years, when catalogues from XMM, Chandra, and SST first appeared.
- About ten papers defied classification, even as "other".

"Hot topic" effects

It has already been noted in connection with Table 2 that cosmology and exoplanet papers are cited more than twice as often as stellar and sun/solar system papers. So wouldn't you think that astronomers working on hot, competitive topics would be eager to get their papers into print (or at least e-print) before someone else scoops them? But no, unless this eagerness in revising is exactly balanced by hot topic referees dragging their e-toes, since Fig. 2 (histograms of delay times of 477 hot topic papers vs. 550 papers on less-cited topics) shows little difference between the two classes.

Correlations with number of authors

There are two results here, one well-known and stable since the 1980s (Abt 1984) and one that rather surprised us. The unsurprising result (Table 3) is that papers with many authors are more cited than papers with one or two, though the correlation is flatter than linear. The smallest author-group that produced a paper with more than 100 citations is seven. Many of these are likely to be self-citations, in the sense that citing and cited papers have at least one author in common. We don't know how many and cannot think of any very amusing way of finding out, or why one should want to. Leave it that your buddies are likely to cite your papers, and conversely. Thus part of the correlation of author numbers with citation numbers is probably a buddy effect. But it is also not unlikely that it takes many people working together to do something difficult and important.

The meaning of "many" has, however, changed over the years, and 30–200 or more names sometimes appear for a large-scale survey or the measurement of a very high energy gamma ray flux from a quasar. Some of these groups have come from the particle accelerator tradition that includes everybody who has helped in any way in the author list. For what it is worth, the medical community has at least partly adopted a custom whereby the people most involved in planning and carrying out a study appear on the first page and the rest (e.g., doctors who contributed only a few patients) as a giant footnote at the end.

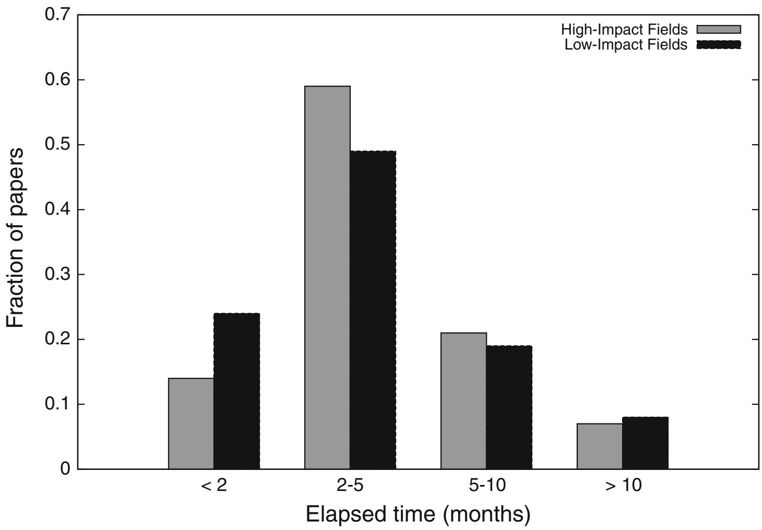

Fig. 2 Histograms of times from submission to acceptance for papers in high-impact fields (cosmology and exoplanets) and low-impact fields (binary stars and sun/solar system). The main point is that they are not very different, apart from the excess of short times for low impact fields (due to the rapid acceptance of many papers about the sun)

Table 3 Numbers of papers, citations, and citations per paper in 3 years after publication for papers with numbers of authors ranging from 1 to more than 100

Number of authors	Citations	Papers	C/P
1	1,362	120	11.4
2	5,174	414	12.5
3	5,991	467	12.8
4	4,170	302	13.8
5	2,943	222	13.2
6	2,173	153	14.2
7	2,066	104	19.9
8	1,228	67	18.3
9	1,227	69	17.8
10	794	45	17.6
11	713	27	26.4
12	607	25	24.3
13	710	30	23.7
14–15	721	28	25.7
16–17	940	29	32.4
18–20	821	23	35.7
21–30	1,572	38	41.4
31–100+	1,096	27	40.6

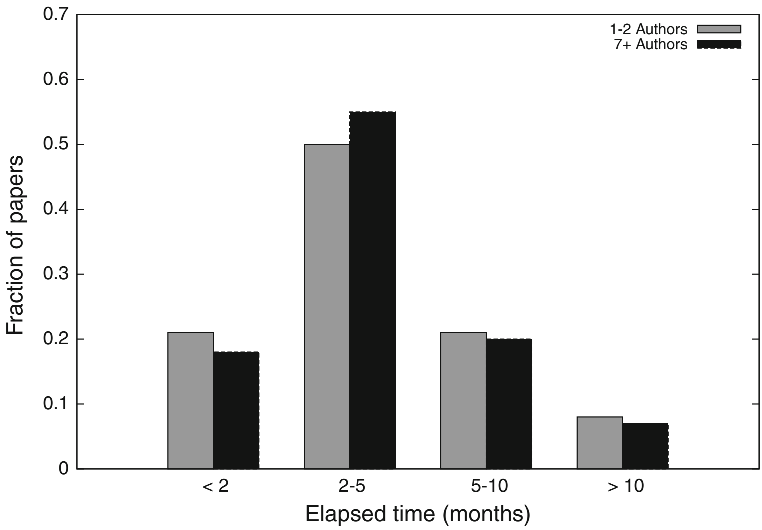

Fig. 3 Histogram of times from submission to acceptance for papers with 1–2 authors versus papers with seven or more authors. The two classes are very similar, a surprise because more time goes into revising than into refereeing, Abt (2009)

The surprising result is that there is no significant correlation of author numbers with time from submission to acceptance. Figure 3 has histograms of lag times for 532 papers with 1–2 authors and 497 papers with seven to more than 100 authors, and they are not very different. The time taken by referees need not depend on numbers of authors (except in so far as many-author papers tend to be slightly longer). But overall more time is taken by authors in revising than by referees in contemplating (something like 55 vs. 44 days in another *ApJ* sample, Abt 2009), and we can only conclude that 1.5 people have as much difficulty as 30 in agreeing with themselves.

Conclusions and suggestions for further work

We have provided a larger, more recent data sample that confirms some things you probably already knew. If you want your papers to be cited frequently, you should work in a rich country, on popular topics, and with lots of co-authors. The effects are factors of 2–3 for each of these variables. Less expected is that the time between when you submit your paper and when it is accepted is nearly uncorrelated with where you work, popularity of topic, and number of authors, and only weakly correlated with numbers of citations received over the next 3 years. The sense of this is that the most-cited papers are accepted on intermediate time scales of two months to a bit more than six. In at least some of these latter aspects, astronomy and astrophysics are very different from stem cell research.

There remains the curious case of the monotonic correlation of citation rate with elapsed time for third world countries. Perhaps for third world authors, who have the fewest citations per paper in our sample, it is more necessary to be published quickly since their research may not have quite as long a "shelf life" as that from more privileged countries.

It would perhaps be interesting to expand the samples in astronomy and stem cell research to include the full range of journals in each, including those with lower impact factors (ratios of citations to papers, roughly) and also to collect numbers for a couple of other disciplines where one can find journals that publish papers from a wide range of subfields written by authors from around the world.

Anyone who would like to try this is more than welcome to what little advice and counsel we have to offer, of which perhaps the most important point is that you have to know the discipline fairly well to identify subfields and figure out who is the submitting author.

Acknowledgments We are grateful to the editors of *New Scientist* for providing more precise numbers for the stem cell analysis than could be gleaned from the graphs in Aldhous (2010a, b) and deeply indebted to Helmut A. Abt for his kind and valuable advice. We also thank our anonymous reviewer for catching several ambiguities that we hope we have addressed.

Appendix

We summarize Table 1 in the following table and figures. Table 4 shows average citations per paper and percent of papers by country (or region) for each of three time bins; for fast (less than 2 months), average (between 2 and 6.31 months), and slow (more than 6.31 months) submission to acceptance times. These time bins correspond to about 20, 60, and 20 % of papers in our sample, respectively. As mentioned in the text, there is no evidence of rapidly accepted papers being more important (in terms of citations) than those that take longer than most (80 %) to be accepted. Figure 4 shows that over all, it is papers in our average time bin that are cited most frequently. Figure 5 shows that the elapsed time is not biased by country (or region) in the astrophysics journal used for this study.

Table 4 Average citations per paper by country/region vs. lag time. Lag time, t, is in months (top row)

	$t < 2$	$2 < t < 6.31$	$t > 6.31$
US	15.6 (0.20)	18.1 (0.62)	15.9 (0.19)
Europe	12.5 (0.21)	19.8 (0.63)	14.9 (0.16)
Other Prosperous	10.5 (0.14)	14.4 (0.69)	11.0 (0.17)
Third World	8.6 (0.19)	8.3 (0.60)	4.8 (0.21)
Total	13.9 (0.19)	17.0 (0.63)	13.9 (0.18)

Numbers in parenthesis are the corresponding fraction of papers for each region's time bin (see Fig. 5)

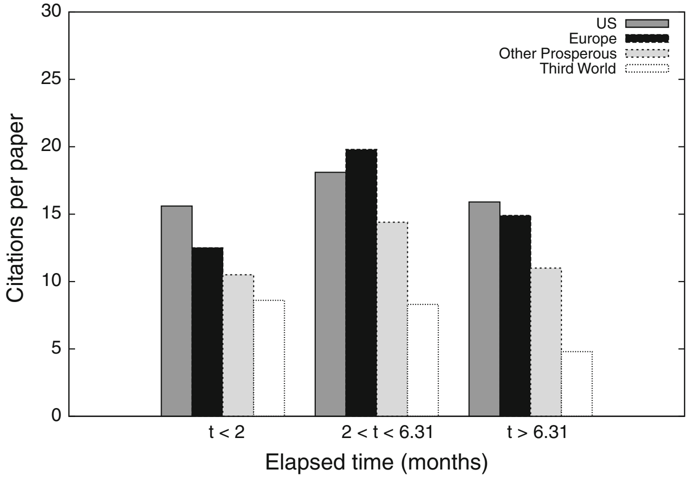

Fig. 4 Average citations per paper by country/region versus time. Citation numbers are for 3 years after publication

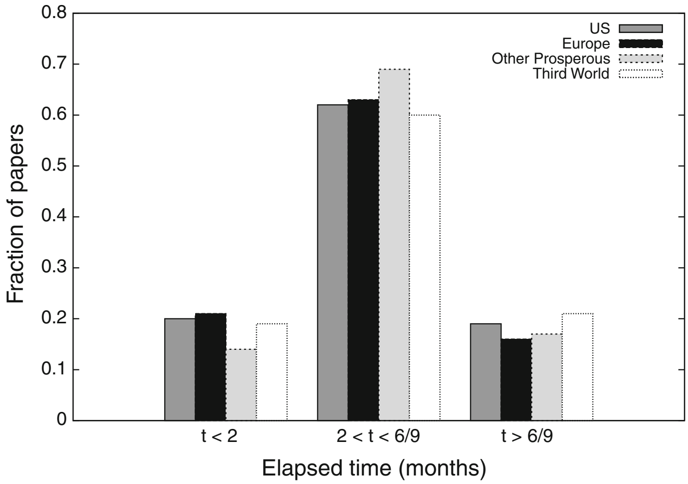

Fig. 5 Fraction of papers by country/region versus time

References

H.A. Abt, Citations to single and multi-authored papers. PASP. **96**, 746 (1984).

H.A. Abt, Reviewing and revision times for The Astrophysical Journal. Public. Astron. Soc. Pac. **121**(11), 1291–1293 (2009).

H.A. Abt, Have we reached a maximum astronomical research output? Publ. Astron. Soc. Pac. **122**(8), 955–958 (2010).

P. Aldhous, The stem cell wars. N. Sci. **12**, 12–14 (2010a).

P. Aldhous, The stem cell wars: data, methods and results. N. Sci. http://www.newscientist.com/article/dn18996 (2010b).

V. Trimble, J.A. Ceja, Productivity and impact of astronomical facilities: three years of publications and citation rates. Astronomische Nachrichten, **329**(6), 632–647 (2008).

V. Trimble, J.A. Ceja, Are American astrophysics papers accepted more quickly than others. Scientometrics, **89**(1), 281–289 (2011).