Citations and Team Sizes

Helmut A. Abt
Kitt Peak National Observatory, NOAO, Box 26732, Tucson, AZ 85726-6732; abt@noao.edu

Received 2016 August 17; accepted 2016 October 7; published 2017 January 12

Abstract

I explore whether small or large teams produce the most important astronomical results, on average, using citation counts as our metric. I present evidence that citation counts indicate the importance of papers. For the 1343 papers published in A&A, ApJ, and MNRAS in 2012 January-February, I considered 4.5 years worth of citations. In each journal, there are larger citation counts for papers from large teams than from small teams by a factor of about 2. To check whether the results from 2012 were unusual, I collected data from 2013 for A&A and found it to be the same as that for 2012. Could the preponderance of papers by large teams be due to self-citations (i.e., citing and cited papers sharing one or more authors)? To answer this, I looked at 136 papers with one to 266 authors and discovered a linear relation that ranges from a 12.7% self-citation rate for single-author papers to a 45.9% self-citation rate for papers with 100 authors. Correcting for these factors is not enough to explain the predominance of the papers with large teams. Then I computed citations per author. While large teams average more citations than small ones by a factor of 2, individuals on small teams average more citations than individuals on large teams by a factor of 6. The papers by large teams often have far more data, but those by small teams tend to discuss basic physical processes.

Key words: sociology of astronomy

1. Introduction

We judge the importance of research papers by the number of citations that refer to them. But are highly cited papers the important ones? That question was addressed in a study titled “Do important papers produce high citation counts?” by Abt (2000). In honor of the centennial of the American Astronomical Society, I asked 53 senior astronomers to select what they thought were the most important papers published in the Astronomical Journal (AJ) and Astrophysical Journal (ApJ) during that century. Then those papers were reproduced in their original formats along with commentaries by the selectors on why those papers were important and how they changed their fields. That material was published in a special large volume of the ApJ (Abt 1999). Then, for those 53 papers as controls I compared their citations counts with the papers published immediately prior and following those papers. It turned out that the 14 selected papers from 1905–1949 received 11.0 ± 3.0 times the citation counts (before 2000) as the control papers and the 37 selected papers from 1950–1974 received 5.06 ± 0.80 times the citation numbers as the controls. I concluded that important papers received much higher citation counts than average papers.

On average, which produces the largest number of citations in astronomy: large teams or small teams of authors? If one wishes to become involved in the most important research, should one join a large or a small team?

I looked at papers published 4.5 years ago in the 2012 January and February issues of Astronomy & Astrophysics (A&A; 326 papers), ApJ (582 papers), and Monthly Notices of the Royal Astronomical Society (MNRAS; 435 papers). Then, for each paper I counted the total citations up to 2016 July. After getting the results, there were other factors involved which are discussed below.

2. The Data

Figures 1–3 show the average annual citations counts for the three journals for 1, 2, 3, etc., numbers of authors (plotted on a logarithmic scale). In each graph the first five points represent papers with 1, 2, 3, 4, and 5 authors. The next four are bins for 6–9, 10–14, 15–30, and 31–500 authors. Horizontal error bars are given for the last two; for the data points of 6–9 and 1–14 authors the error bars are too small to show outside the data points. The three distributions are not statistically different from each other. For the three journals, the ratio of annual citations for papers with 10 or more authors divided by that for papers with less than 10 authors is 2.26 ± 0.17. Therefore, large teams write papers producing about twice as many citations as small teams.
Could it be that papers in 2012 are not typical? To check for this, I collected the same statistics for 306 A&A papers published in 2013 January and February. Figure 4 shows the annual citations for papers with different numbers of authors. It is statistically the same as Figure 1. Therefore, the results for 2012 do not seem to be unusual.

However, there are two reservations. One concerns self-citations. Many self-citations are justified because often each paper is a step in the progressive development of a field, so one paper builds on the previous ones. However, it is obvious that...
if a paper has many authors, they can build on their previous papers, causing many citations, so I explored the occurrence of self-citations. A self-citation is defined as one in which the citing paper and the cited paper have at least one author in common. Searches for self-citations are tedious. An extreme example is the 242-author paper by Abbasi et al. 2012 (“Observation of Anisotropy in the Galactic Cosmic-Ray Arrival Direction at 400 TeV with IceCube”); was it cited by the 566-author ApJ paper Aab2012 (“Searches for Large-scale Anistropy in the Arrival Directions of Cosmic Rays Detected above Energy of 10^{19} eV at the Pierre Auger Observatory and the Telescope Array”)?

I looked at 14 single-author papers in the 2012 ApJ, 45 two-author papers, 46 four-author papers, 21 papers with seven or eight authors, six papers with 38–67 authors, and three papers with 261–266 authors. Figure 5 shows the results; it plots the self-citation frequency against the log of the number of authors. The linear least-squares fit is 12.7% + 16.6% log A. That means that 12.7% of the single-author papers produce self-citations, while 45.9% of the 100-author papers produce self-citations.

Are these self-citation frequencies enough to explain the increase in mean citation numbers with author sizes? For single-author papers, we should delete 12.7% of the citations and for 100-author papers we should delete 45.9% of the citations. After we subtract all the self-citation numbers from Figures 1–3, the large teams still produce more citations than the small teams; self-citations do not change that.

In a multi-author project, one person does not do all of the work. Therefore, no one person deserves all of the credit (citation counts) for the project. We are rarely told of the distribution of planning, work, and interpretation among the authors. Often, but not always, the most active authors are listed first and the remainder are listed alphabetically after, although it rarely seems to be the case that most or all of the credit should be given to one person. Therefore, our second reservation concerns the distribution of credit given to each author. For statistical purposes, we will assume equal credit for each author, so that the total citation counts should be distributed equally among the authors. The ADS understands that and lists “normalized citation counts”, i.e., the totals divided by the numbers of authors.

When I divided the total citations counts by the number of authors for each paper, I obtained the distribution of normalized citation counts for the 582 2012 January and February ApJ papers that is shown in Figure 6. Here it appears that per person, the individuals on small teams produce far more citations than individuals on large teams. In fact, papers with one or two authors produce six times as many citations per person than papers with >10 authors per paper.

This may seem to be a surprising result because, for instance, a large team can generally collect much more data than a small one. So what is it about some of the papers by small teams that produced such a large number of citations per person? Table 1 lists the top six papers according to normalized citations as plotted in Figure 6. These papers can be categorized as fundamental studies, rather than data-heavy papers. In contrast, Table 2 lists the top eight papers by team size. They are
primarily data-rich papers. Therefore, small teams tend to work on basic physical processes while large teams tend to produce data-rich papers.

I conclude that large teams produce more citations than small ones, but that individual authors on small teams produce more citations per author than those on large teams.

This research has made use of NASA’s Astrophysics Data System.

References
Aab, A., Abreu, P., Aglietta, M., et al. 2012, ApJ, 794, 172
Abassi, R., Abdou, Y., Abu-Zayyad, T., et al. 2012a, ApJ, 744, 1
Abassi, R., Abdou, Y., Abu-Zayyad, T., et al. 2012b, ApJ, 745, 45
Abassi, R., Abdou, Y., Abu-Zayyad, T., et al. 2012c, ApJ, 746, 33
Abramowski, A., Acero, F., Aharonian, F., et al. 2012a, A&A, 537, A114
Abramowski, A., Acero, F., Aharonian, F., et al. 2012b, ApJ, 746, 151
Abt, H. A. 1999, ApJ, 525, 1C Part
Abt, H. A. 2000, Scim, 48, 65
Ade, Aghanim, N., Arnaud, M., et al. 2013, A&A, 550, A134
Ajello, M., Allafort, A., Baldini, L., et al. 2012, ApJ, 744, 80
Aleksić, J., Alvarez, E. A., Antonelli, L. A., et al. 2012, ApJ, 746, 80
Ekström, S., Georgy, C., Eggenberger, P., et al. 2012, A&A, 537, A146
Graham, A. W. 2012, ApJ, 746, 113
Haardt, F., & Madau 2012: “Radiative Transfer in a Clumpy universe. IV New Synthesis Models of the Cosmic UV/X-Ray Background”
Kraus & Ireland 2012: “LkCa 15: A Young Exoplanet Caught at Formation?”
Leitner 2012: “On the Last 10 Billion Years of Stellar Mass Growth in Star-forming Galaxies”
Graham 2012: “Breaking the Law: The Mbh–M_spheroid Relations for Core-Sérsic and Sérsic Galaxies”
Leinhardt & Stewart 2012: “Collisions between Gravity-dominated Bodies. I. Outcome Regimes and Scaling Laws”
Aab, A., Abreu, P., Aglietta, M., et al. 2012, ApJ, 794, 172
Abassi, R., Abdou, Y., Abu-Zayyad, T., et al. 2012a, ApJ, 744, 1
Abassi, R., Abdou, Y., Abu-Zayyad, T., et al. 2012b, ApJ, 745, 45
Abassi, R., Abdou, Y., Abu-Zayyad, T., et al. 2012c, ApJ, 746, 33

Table 1
Early 2012 ApJ Papers with the Top Citation-collectors per Person

| Citations per Person | Reference and Title |
|----------------------|---------------------|
| 130.0                | Haardt & Madau 2012: “Radiative Transfer in a Clumpy universe. IV New Synthesis Models of the Cosmic UV/X-Ray Background” |
| 105.5                | Metzger & Berger 2012: “What is the Most Promising Electromagnetic Counterpart of a Neutron Star Binary Merger?” |
| 88.5                 | Kraus & Ireland 2012: “LkCa 15: A Young Exoplanet Caught at Formation?” |
| 76.0                 | Leitner 2012: “On the Last 10 Billion Years of Stellar Mass Growth in Star-forming Galaxies” |
| 76.0                 | Graham 2012: “Breaking the Law: The Mbh–M_spheroid Relations for Core-Sérsic and Sérsic Galaxies” |
| 56.5                 | Leinhardt & Stewart 2012: “Collisions between Gravity-dominated Bodies. I. Outcome Regimes and Scaling Laws” |

Table 2
Early 2012 ApJ Papers with the Largest Numbers of Authors

| Authors | Reference and Title |
|---------|---------------------|
| 447     | Abramowski et al. 2012b: “The 2010 Very High Energy γ-Ray Flare and 10 Years of Multi-wavelength Observations of M87” |
| 266     | Abbasi et al. 2012b: “Neutrino Analysis of the 2010 September Crab Nebula Flare and Time-integrated Constraints of Neutrino Emission from the Crab Using IceCube” |
| 265     | Abbasi et al. 2012a: “Time-dependent Searches for Point Sources of Neutrinos with...IceCube” |
| 261     | Abbasi et al. 2012c: “Observations of Anisotropy in the Galactic Cosmic-Ray Arrival Directors at 400 TeV with IceCube” |
| 203     | Abramowski et al. 2012a: Discovery of hard-spectrum γ-ray emission from the BL Lacertae object 1ES 0414+009 |
| 158     | Shin et al. 2012: “Micro lensing Binaries Discovered through High-magnification Channel” |
| 154     | Aleksić et al. 2012: “Detection of the γ-Ray Binary LS +61°303 in a Low-flux State at Very High Energy γ-Rays with the MAGIC Telescopes in 2009” |
| 149     | Ajello et al. 2012: “Fermi Large Area Telescope Observations of the Supernova Remnant G8.7-0.1” |