

Information Use in Astronomy

William P. Claspy

Case Western Reserve University, Cleveland, Ohio 44106, USA

Abstract. This study seeks to examine astronomers' use of their literature through a citation analysis of a very current sub-section of scientific literature in the field. The paper specifically analyzes currency of references, types of sources, format of sources, and cross-disciplinary sources. Research published in refereed journals is the most frequently cited material among astronomers, making up nearly 80% of the total number of references. Of these, the majority are references to literature from the core journals in the field. Also, currency is of relative importance to researchers in the field, as nearly 56% of all references were to materials published in the last seven years, and nearly 90% to that published since 1980. Only 4% were to materials published before 1970.

1. Introduction

Research in the field of astronomy is, like research in all scientific fields, dependent on the record of previous work on which it can then build. Astronomy libraries are relied upon to provide and maintain collections of this published information so that scientists can make reference to the work that has already been done in their field. It is the astronomers' use of their literature that this study seeks to analyze.

Many fields of research, especially those in the sciences, have begun to use the electronic transmission of the results of research. Such transmission greatly speeds the results of researchers to others in the field. Astronomy, in particular, has historically been among the first disciplines to embrace electronic technologies, in part due to the large amounts of data produced by research in the field. Computerization of these data simply makes them easier to access and to use. With the advent and growth of electronic protocols such as FTP and the World Wide Web (WWW), it has become easier to electronically access research, both published and unpublished.

1.1. Purpose of the Study

The purpose of this study is to determine how researchers in astronomy are using published and unpublished information in their work. Some of the questions this study asks are: how frequently do researchers in the field make reference to older published materials? To what extent do researchers make use of cross-disciplinary materials? To what extent are unpublished materials used?

Another trend that this paper seeks to study is what appears to be a growing use and reliance upon electronic transmission of scholarly information. As of the writing of this paper, the core journals studied are becoming available in full text electronically via the WWW. However, preprints in astronomy have been available electronically via FTP for many years. Preprint servers such as the one at Los Alamos National Laboratory (<http://xxx.lanl.gov/index/astro-ph>) continue to be very popular. This study seeks in part to see to what extent these electronically disseminated publications are becoming cited references.

Cross disciplinary references are also included in this research, in the hopes of finding out how astronomers use literature published outside their field. This information can be useful in making campus-wide collection decisions regarding issues such as journal cancellations or collection (de-)centralization.

The results will hopefully give an indication as to the changing trends of how scientists in the astronomy community are using the resources within, as well as those without, our libraries. They can help guide us in making collection development and management decisions, as well as the increasingly difficult resource allocation decisions with which we are faced.

2. Literature Review

The bulk of the relevant previous studies were found within the scientific literature, written by astronomers. One astronomer in particular has shown a distinct interest in the literature of astronomy: Helmut A. Abt of Kitt Peak National Observatory. He has written many articles focusing on various bibliometric aspects of the literature of astronomy. Beginning in 1981 Abt has published a series of articles each of which examines a particular aspect of the literature of astronomy.

In an interesting half-life study Abt (1996) studies the citations, from 1955 to 1994, to the 165 papers published in the *Astrophysical Journal* and *Astrophysical Journal Supplement* in 1954. He finds that the number of citations shows an exponential decay with a half-life of 29 years. Further, he finds that papers receiving more citations than others during the first five years do not have longer half-lives. In order to account for the rapid growth of the field since the 1950s he then normalizes the figures to a constant literature output, where the average half-life becomes six years, a figure much closer to expected literature half-lives.

In the most interesting of Abt's papers (Abt 1995), at least in relation to the goals of this research study, he looks at the "changing sources of published information". In his study, he examines the types of references in papers published during the first month of 1972, 1982, and 1992 in two major journals in the field. He found that, somewhat predictably, reference rates to in-house technical reports and observatory publications fell over the period (from 12% to 1%) and references to conference proceedings increased over the period (from 1% to 10%). This trend is predictable because fewer observatories and departments of astronomy are distributing their preprints as was commonly done in the past. Also, publication of conference proceedings has become more common with the growth of the International Astronomical Union *Symposia* and *Colloquia* and the *Conference Series* of the Astronomical Society of the Pacific.

Abt was surprised to find, however, that there was a low rate (less than 3%) of references to preprints, showing that these types of publications were not replacing journal papers (78%) as the primary source for reliable information. He found a similarly low reference rate to theses and review papers, and a slightly declining (from 6% to 4%) rate of references to monographs.

The increase in references to conference proceedings was further researched by Vagiswari & Louis (1995). They analyzed the change in referencing patterns among the astronomers, and in particular looked at the increase in the number of conference proceedings published and referenced since 1980. They found that the number of proceedings published rose from 131 in 1980 to 267 in 1990. Studying references in the *Astrophysical Journal*, the authors found that the number of references to these proceedings rose similarly, from 1.48 references per paper in 1980 to 3.8 in 1990, an increase of 150%. The growth in importance to researchers of these published proceedings has a large impact on library budgets: we now need to identify funds to pay for these new sources of information used by astronomers.

Similarly, Bouton & Stevens-Rayburn (1995), in a discussion of their libraries' preprint databases, discuss the growth in the number of preprints issued. They reported that the National Radio Astronomy Observatory library database contained 3207 preprints in 1988, 1602 of which were unpublished. In 1995, the database contained 6315 preprints, 1604 of these unpublished. The growth trend in preprints, both print and electronic, is also addressed by this research paper.

Bibliometric studies similar to Abt's (1995) have been performed recently using NASA's Astrophysical Data System (ADS), a World Wide Web based database which indexes, abstracts, and selectively makes available in full text an astonishing amount of the published literature in astronomy. For example, Schulman et al. (1996) use the ADS database to study trends in astronomical publication similar to Abt's analyses. By using computer scripts, the ADS database can easily be queried to study the fraction of papers with one author or the average length of papers published, for example.

Through an arrangement with the Institute for Scientific Information, the producers of Science Citation Index, ADS has recently included listings of references for papers in the database, making reference tracking via the database quite simple. However, ADS has included those references that also exist within the ADS database, so that a hypertext link can be made to the cited article. So, for example, if an author makes a reference to a paper written in 1950, it will not appear in the references listed on ADS since the database only goes back to the early 1980s.

3. Methodology

In order to gain a clear picture of the current use of literature by researchers in astronomy, several criteria were set to perform the citation analysis outlined in this paper. An overall view of the researcher's current needs was aimed at, so the decision of which journals to inspect was made as follows.

A group of core journals was chosen for the study. This group is widely and uniformly considered the group of standard journals of research in astronomy,

and represents the majority of important research performed in the field. Other journals certainly exist, but represent either specialized fields of research within astronomy, contain largely public-interest articles (i.e. not research), or are review journals.

The journals chosen are the *Astronomical Journal* (AJ), *Astronomy and Astrophysics* (A&A), *Astronomy and Astrophysics Supplement Series*, *Astrophysical Journal* (ApJ), *Astrophysical Journal Supplement*, and *Monthly Notices of the Royal Astronomical Society* (MNRAS). While these journals represent the core journals in the field, they also represent a fair diversity of authors. AJ and ApJ are edited and published in the United States, A&A in Europe and MNRAS in Britain. They also all publish papers from the full spectrum of subjects within the field of astronomy, including the supplement series of ApJ and A&A which are most often longer papers containing data intensive work such as catalogs or other large data sets. Studied were citations from letters (which are generally shorter articles or brief reports that are published more quickly than a full article might be) as well as papers from these journals.

Performing a computer assisted database search of ADS was considered as a method of obtaining data for this research paper, however it was found unsuitable for finding the type of citation analysis information that most closely fit the purpose of this study. Links in the ADS system do not include references to works produced before the early 1980s. The use of references to these works is part of what this study seeks to examine. Hence, it was decided that performing the analysis on ADS would not be a suitable method for this study.

In some ways, the methodology of this study is similar to those studies conducted by Abt (1995) and seeks to replicate his studies with a different collection of data to see if his conclusions are still valid. One difference in this study is that it examines the role of electronically published information and what impact it has had overall in the published literature of astronomy. Of course, it is difficult to gauge this simply from examining references in papers. In an article written in 1982, for example, the probability that the author electronically obtained a preprint that he cites is fairly low. In 1997, however, that likelihood becomes very high indeed.

Similarly, authors now rely heavily upon electronic mail to communicate. Such ease of communication makes the production of the papers with more than two or three authors a much easier task than it would have been ten or fifteen years ago. Indeed, electronic communication may be one very strong cause for the recent increase in the number of papers with ten or more authors, the possibility of which was nearly impossible before the advent of electronic mail.

Even more difficult to ascertain in a study such as this is the impact of electronically disseminated refereed journals. Today, four of the seven journals studied here are available full text electronically via the WWW. Full text of recent articles from all three are available via ADS. When the author makes reference to an article published in a 1996 issue of the ApJ, did he obtain and read that copy in paper (either through his own subscription or from his library's copy) or did he download the article in full text from the WWW version, or via the ADS? This kind of question is more difficult to answer, as it is not necessary for the author to designate in his reference in which format he read the article.

Authors do cite other types of electronically formatted materials. For instance, preprints obtained via one of the preprint servers are noted as such, as are any other types of publications that exist only in electronic format, be it residing on an FTP server or on the WWW. These references in particular are categorized separately in this study.

Also of interest was the number of references to literature outside the field. This study counts the number of citations to information sources from other scientific specialties such as physics and chemistry. Such information can be quite valuable when making collection development decisions, especially those involving consortia, resource sharing, and cooperative collection development, or even the purchase of access rights to electronic journals. If a campus wide decision needs to be made concerning the purchase of such rights, cross disciplinary use figures can be used to justify the purchase. These data can also be quite useful in studies related to centralizing or decentralizing collections.

Since it is not uncommon for one issue of each of these journals to contain forty or more papers, each with forty or more citations, defining which papers are to be included is not simple. The most practical solution is that several issues of each title should be consulted, and a random sample of papers from these issues included in the study. This would isolate those issues that focus on single issues, as well as take into account issues that are arranged topically. Issues from the journals noted above were selected from those published from October to December 1996. Ten articles were randomly selected from each journal. The total number of papers examined was 60, and that of citations analyzed was 1735.

The aspects that seemed most important to tally within the focus of this paper were as follows: (i) type of publication cited. This included journal article, preprint, observatory publication, conference proceeding, monograph, personal correspondence, dissertation or thesis and "in preparation", (ii) of journals cited, reference to the core journals was counted, (iii) cross disciplinary journals, (iv) year of publication of cited item, (v) format of cited item: either print or electronic.

The information within these categories was recorded onto a coding form and then entered into the statistical computer package SAS for compilation. The data was analyzed in several ways. The figures calculated were percentages of references by type of information (journal article, monograph, etc.), percentage of total references that refer to core journals, percentage that refer to journals from outside the discipline, percentage of references to information in electronic format. Also computed was a frequency distribution of number of references by year of publication (arranged as <5 years, 5–10 years ago, 11–15 years ago, etc.).

4. Analysis of Data

After coding the references and entering them into the SAS package, frequencies and percentages for each category were computed. The first category counted was type of publication (Table 1). Not surprisingly, the most frequently referenced materials were from the journal literature at 79.4%. This figure corresponds quite closely with that calculated by Abt (1995) in his study (78%). Of those references that were to journal literature, 65.8% referred to articles from the core journals. This figure represents a significant proportion of the literature

cited. Obviously, the core journals in astronomy present a respected and highly used source of information for researchers.

The next most frequently cited type of material were both monographs and conference proceedings. These are very similar types of materials, as conference proceedings are commonly published in book format. Combined, these materials accounted for 12.5% of the references studied. And, of course, monographs often play a different role in the astronomy community than supporting research: that of curriculum support in academic settings.

Table 1. Distribution by Publication Type

Publication Type	Frequency	Percentage
Journal	1378	79.4
Monographs	124	7.1
Conference proceedings	93	5.4
Preprints	51	2.9
Observatory publications	32	1.8
In Preparation	25	1.4
Dissertations/Theses	23	1.3
Personal correspondence	9	0.5
Total	1735	100.0

The other material types accounted for only small percentages of the total. Observatory publications accounted for only 1.8% of the total. As noted earlier, this figure is not surprising, considering the small number of observatories worldwide that continue to publish research findings on their own. Dissertations and “in preparation” materials accounted for 1.4% and 1.3% respectively and personal correspondence for less than 1%.

There was an equally small percentage of the total references to preprints: only 2.9% referred to this type of publication. This figure also matches quite well with the figure calculated by Abt. His contention that preprints have not replaced journal literature as the main source for reliable information is a strong one. Even with the growth of rapidly available preprints via FTP and WWW servers, researchers continue to rely on refereed journal literature to support their research information needs.

Astronomers are an insular group: 92% of the references, regardless of type of publication, were to some type of astronomical literature. Of the subject areas other than astronomy, literature in the field of physics was most often cited, by 4.7% of the references. After physics came those sources that are broad in their coverage, journals such as *Science* and *Nature* which publish scholarly articles in all fields of the sciences.

Although a majority of the references studied were to materials published since 1970, a surprising amount of the references were to items published before 1970. But currency seems to be an extremely important feature to researchers, as ~56% of the references were to materials published since 1990 (see Table 2).

Very few of the references specifically noted an electronic source. Less than 1% of the references were to electronic material. These references were to one of two types of sources: preprints obtained from the Los Alamos preprint server or either manuals or handbooks published on the WWW. The number can be somewhat misleading, however, in that reference to a journal article obtained electronically need not specify the method by which the article was obtained. A

reference to an article published in the ApJ looks exactly the same regardless of what format the author read it. In this regard, a citation analysis was not the best research method to employ. For discovering the extent of the use of electronic sources in astronomy research, other methods should be employed.

Table 2. Distribution by Publication Date

Publication Date	Frequency	Percentage
1995–96	311	17.9
1990–94	659	38.0
1980–89	529	30.5
1970–79	169	9.7
1950–69	52	3.0
1901–49	15	0.9
Total	1735	100.0

5. Conclusions

The results of this study show several important trends in the use of information sources by astronomers. The references that they use in their research is a mirror of the sources they use, and as such, reflect those library materials that are of highest importance in a library research collection.

Clearly, based on the first set of data collected here as reported in Table 1, research presented in the journal literature is of highest importance to researchers. Astronomers are concerned that the science used to support their own findings has been rigorously examined, not only by the authors of the research, but also by the traditional series of referees and editors involved in peer reviewed journals. A large percentage of the journal titles cited were refereed journals. This type of research publication is the most authoritative.

Conference proceedings have indeed become a more popular forum for publishing research findings, but most likely will not gain on journals for the most cited references. This is due to several reasons. While interesting findings are reported at conferences, the papers submitted are rarely refereed. Usually, these papers are short and do not give the full data collected in the research or a detailed analysis, but rather offer a quick overview of the author's findings. Often, the author will subsequently publish the paper in a peer-reviewed journal. According to Abt (1995), referees recommend that, whenever possible, references to papers presented at conferences are replaced by citations to journal papers giving the full data, when such a paper is eventually published.

Despite the relatively small percentage of references to preprints, it must be noted that this type of publication is an important one for astronomers. In gathering information while preparing studies, preprints and electronic preprints are invaluable tools for scientists to have the most complete knowledge of current trends in their field. More often than not, references to preprints become references to resulting published papers during the course of refereeing.

From a librarian's point of view, this reliance on the journal literature is important to note. Journal subscriptions, be they in paper format or electronic format, continue to be of highest importance to researchers in the field.

Another clear trend identified by this research is that the astronomy community largely regards the papers published in their core journals as the most reliable form of reference. These core journals are the publishers of the most reliable research findings in the field. Such a preponderance of the references are published in these journals that their subscriptions, as well as subscriptions to indexing tools to access the material within the journals, must be maintained by all libraries whose aim is to support research in astronomy.

Purchase of monographic material also needs to be maintained. When taken together, monographs and conference publications (often published in monographic format) make up a relatively large subset of the references made within the literature. Since increasingly larger percentages of budgets are being taken up by the maintenance of journal collections, the choices made when purchasing monographs are critical. Care must be taken to select those monographic titles and conference publications of most use to the local community.

Although this study did not find a large number of references to explicitly electronically formatted materials, the selection and maintenance of these types of publications should also be considered as a primary goal of astronomy libraries. Clearly, research findings available electronically will be of more use to scientists than more traditional formats, if based only on convenience matters. Other considerations in favor of this format should be considered as well, including increased accessibility via electronic indexes and savings in costs of physical storage for traditional printed materials.

Finally, further research on the information use of astronomers is warranted. While this study and others make clear the overall trends of the types of materials used by scientists in the field, other studies could further our understanding of the dynamics of their research methods. An understanding of the impact of electronically available material would be of tremendous use to librarians who face tough decisions on the issue. Also, research could be conducted on the use of monographic materials in our libraries, based on use both in support of research as well as support of curriculum (for those libraries associated with teaching university programs).

Acknowledgments. I would like to thank the Friends of LISA III and Case Western Reserve University for making it possible for me to attend this meeting.

References

- Abt, H. A. 1995, *PASP*, 107, 401
Abt, H. A. 1996, *PASP*, 108, 1059
Bouton, E. N., & Stevens-Rayburn, S. 1995, *Vistas Astronomy* 39, 149
Schulman, E., Powell, A. L., French, J. C., Eichhorn, G., Kurtz, M. J. & Murray, S. S. 1996, *BAAS* 189, #06.08
Vagiswari, A., & Louis, C. 1995, *Vistas Astronomy* 39, 173