

我校学报论文的引文初步分析

刘吉卿

(图书馆)

摘 要

本文以福州大学学报自然科学版1978年复刊以来至1982年第一期的全部自然科学论文为样本,统计了各篇论文引用文献的一些标志和特征,得出我校学报自然科学论文的引文的一些初步结论:引文的数量范围为0—18篇,平均6篇。引文中英语的参考文献最多,占总引文量的44.41%,占外文总引文量的76.45%。大部分情报源来自英语期刊和图书。引文出现了引用产品样本及说明书、学位论文、教材讲义及内部资料,体现了我校是理工科大学教学与科研、理论与实践相结合的特征。近二十年来所发表的科学论文是主要的引用对象,占总引文量的78%。本文统计的结果可靠性较好。

一、引 言

科学论文中所列的参考文献称为论文的引文。科学工作者在反映他们劳动成果的科学论文中,一方面为了尊重前人或他人的劳动,另一方面为了阐述和论证自己观点的正确性,往往在论文文末列举被引用的参考文献,这就是说,作者在整个科研活动的过程中,往往需要参阅、消化、吸收和借鉴这些文献的知识。因此,通过对这些引文的统计、分析和评价,可以了解读者的阅读倾向,侧面反映出文献知识被利用的情况,充分发挥这些文献的作用和提高图书情报工作的服务质量都是有益的、有意义的。

我校是理工科大学,肩负培养科技人才和开展科学研究双重任务。我校图书馆藏书以自然科学为主,服务对象主要是教师、学生和工程技术人员,满足教学和科研的书刊需要为目的。我校学报的任务是^[1]坚持四项基本原则,贯彻“双百方针”,报导科研成果,开展学术交流,推广先进经验,鼓励发明创造,调动一切积极因素推动教学、科研和科学实验活动的深入开展。因此,学报论文在一定程度上可以反映我校教学、科研、学术水平和实践经验等方面的成果。本文是试图通过对学报论文的引文分析这一客观实际的侧面,从中进行分析与评价,为我校图书情报的供求工作提供参考依据,使它在教学、科研和科学实验中能够充分发挥作用并提高图书情报工作的服务质量。

本文1982年10月18日收到。

二、分析方法

本文统计与评价的对象是我校学报自然科学版78年复刊以来至82年第一期共8期的全部自然科学论文123篇(包括研究简报和综论)的引用参考文献804篇(引文总数)。这就是说凡在我校学报78年复刊以来各期上发表的123篇论文(不管其有无引文)都作为本文统计的对象,我们之所以这样做,其目的是能够更具体地反映学报复刊以来804篇论文的引文整体面貌和某些特征与规律。

本文参考了王崇德对论文的统计和评价,^[2]并运用引文分析的方法,把上述学报论文分别按篇为单位,统计它们的引文数量、语种、类型和出版年度等,然后把这些统计数据加以归纳、整理和描述,最后进行分析讨论作出初步评价。

三、结果与讨论

1. 引文平均值

在本文统计的123篇论文中,引文最多者为100篇(系一篇综述性论文的引文),最少者为0篇(即没有列出引文)见表1。

表1 引文量在论文中的分布

引文量 (篇)	论文频数 (篇)	引文量× 论文频数 (篇)	引文量 (篇)	论文频数 (篇)	引文量× 论文频数 (篇)	引文量 (篇)	论文频数 (篇)	引文量× 论文频数 (篇)
0	14	0	8	2	16	16	1	16
1	4	4	9	4	36	17	0	0
2	11	22	10	8	80	18	0	0
3	24	72	11	1	11	19	1	19
4	13	52	12	4	48	20—74	0	0
5	9	45	13	0	0	75	1	75
6	11	66	14	4	56	76—99	0	0
7	8	56	15	2	30	100	1	100
合计							123	804

从表1可以看出,绝大多数的论文引文量均在0—19之间,20—100之间的引文量仅有75及100的各出现一次,在科学论文中出现大数量的引文是比较少见的。如果把表1所列的引文量0—19之间变动的数据系列,规定组距为3的话,可以把它们划分为7组,用统计学方法计算引文的算术平均数M,计算结果列于表2。

表2 引文分组分配表

序号	分组区间	组中点 A	频数 F	F·A
1	0—2	1	29	29
2	3—5	4	46	184
3	6—8	7	21	147
4	9—11	10	13	130
5	12—14	13	8	104
6	15—17	16	3	48
7	18—20	19	1	19
合计			121*	Σ661

引文算术平均值 $M = \Sigma F \cdot A / N = 661 / 121 = 5.46$

(*亦即为样本的容量)

从表2所列数据的计算结果，论文引文的算术平均值M为5.46，这数值仅反映所统计的论文数量平均概念，并不说明论文的引文相互之间在数量上的波动程度及变化范围。对于这些变化的数据系列，尚需进一步进行统计学的运算后加以描述。用于统计学计算的基本数据列于表3。

表3 用于统计学计算的基本数据

序号	V	X	X ²	序号	V	X	X ²
1	4	-1.46	2.1316	9	16	+10.54	111.0916
2	9	+3.54	12.6319	10	8	+2.54	6.4516
3	2	-3.46	11.9716
4	5	-0.46	0.2116	117	6	+0.54	0.2916
5	3	-2.46	6.0516	118	11	+5.54	30.6918
6	10	+4.54	20.6116	119	8	+2.54	6.4516
7	1	-4.46	19.8916	120	5	-0.46	0.2115
8	7	+1.54	2.3716	121	4	-1.46	2.1316
					Σ629	±84.70	Σ2023.5236

表中：V——每篇论文的引文数量

X——每篇论文的引文量与算术平均值的偏差

X²——偏差的平方

根据表3所列数据，我们可以作一些计算。

(1) 标准误差σ——这是重要的统计参数之一。它表明数据系列平均变化程度的一个指标，计算公式如下：

$$\sigma = \pm \sqrt{\frac{\sum X^2}{N}} = \sqrt{\frac{2023.5236}{121}} = 4.1$$

标准误差可以判断是否属于我们已知的观测数据系列。例如拟考察引文量为18的论文在学报论文中占有的比例，因为18可以近似地写作 $5.46 + 3 \times 4.1 = 17.76$ ，概率论已经证明：99.7%的数值是在 $M \pm 3\sigma$ 范围之内。所以可推断附有18篇以上引文的论文，在上述统计的学报论文中不大于0.3%，这就是说学报论文的引文均在0—18篇之间。

(2) 算术平均误差 m ——可按下式计算：

$$m = \pm \frac{\sigma}{\sqrt{N}} = \pm \frac{4.1}{\sqrt{121}} = 0.37$$

m 是具有以局部的算术平均值来推测全体的算术平均值特征的数值。已知 M 及 m ，就可以判断 M 的可靠性。与 σ 相似，概率论已证明：当同样性质的观测数据相当多的时候，有99.7%不超过 $M \pm 3m$ 的范围。这就是说如果统计更多篇学报论文的引文时，其平均值 M 有99.7%的可能波动在4.35—6.57之间。（由 $5.46 \pm 3 \times 0.37$ 计算而得）

(3) 变化系数 V_1 ——可按下式计算：

$$V_1 = \pm \frac{100\sigma}{M} = \pm \frac{100 \times 4.1}{5.46} = 75.1\%$$

这是表示比较两个或两个以上算术平均值变化程度特征的数值。

(4) 精确度 P ——这是说明观察可靠的数值，它愈小愈可靠。 P 值可按下式计算：

$$P = \pm \frac{100m}{M} = \pm \frac{100 \times 0.37}{5.46} = \pm 6.7\%$$

一般认为 P 值应在 $\pm 4\%$ — $\pm 6\%$ 才能充分保证观测与统计的可靠性。我们计算的 P 值略高于上述范围，说明统计结果的可靠性略低。

从以上统计计算结果可以初步看出，我校学报复刊以来论文平均引文量为近6篇，这数值是偏低的。目前我国化学基础研究论文平均引文量为12篇^[2]，国际上表面化学研究论文的引文平均数为18篇^[2]。我校学报论文的引文平均数偏低的原因，我们认为是一校科研工作中情报吸收量偏低所致。

2. 引文的语种

本文统计了123篇论文引文的语种，其结果见表4图1。

从图1可以看出论文作者引用参考文献有6种语种，其中英语和中文所占的比重最大，分别为44.41%和41.92%，而英语在外文总引文量中占76.45%（见表4），其中英语期刊和图书占英语总引文量92.72%，这说明如果一位从事科技工作者，只要他能熟练地掌握英语，加上中文就可以获得86.33%的情报源。从表4统计数字来看，我们还应该注意到，没有中文专利文献可供引用。究其原因可能是我国专利法尚未颁布实施所致，这是符合我国当前实际情况的。但是从科技报告、内部资料和学术会议论文集的中文引文的比重来看显然是较大的（总的比重仍偏低），这在一定程度上弥补了没有中文专利文献的欠缺。学术会议论文集、科技报告、专利说明书和内部资料等第一手外文文献的引用数量也是较少的，其原因是我国科研工作中情报的吸收量偏低。这说明我们在提供和使用情报两方面都有不足之处，

表4 引文语种的比较

语种	类型		期刊	图书	图	专利说明书	科技报告	学位论文	学术会议论文集	内部资料	文摘	产品样本及说明书	手册	综述	教材讲义	总计	占外文引量的%	占总引文量的%
	类	型																
外语	英语		263	68		9	1		2	1	10	1	2			357	76.45	44.41
	俄语		48	32						1						81	17.35	10.07
	日语		11	1												12	2.57	1.49
	德语		12			4										16	3.42	1.99
	法语		1													1	0.21	0.12
中文	总计		335	101		13	1		2	2	10	1	2			467	(100)	
	占外文引量的%		71.73	21.63		2.78	0.22		0.43	0.43	2.14	0.21	0.43					
	中文		132	12.3			12	2	9	42	2	.3	6	1	5	337		41.92
占中文引量的%		39.17	36.50			3.56	0.59	2.67	12.46	0.59	0.89	1.78	0.30	1.48	(100)			

有待今后改进和提高。我们应该考虑是否需要扩大进口期刊和图书、收集和交换中外文第一手参考文献，以满足科研的需要，这是情报图书工作者与情报使用者双方共同的责任。此外，我们还可以看到学位论文、教材讲义、产品样本及说明书等类型中文文献，不同程度地在被引用，这体现了理论研究和应用研究、教学和科研相结合的特征，和我校属于理工科性质的大学也是符合的。

3. 引文的类型

科学研究工作经常参阅并引用的科学文献，主要有以下15个类型：

一次文献：①期刊、②图书、③专利说明书、④科技报告、⑤学位论文或毕业论文、⑥学术会议论文集、⑦内部资料、⑧口头交换及私人通信材料。

二次文献：⑨文摘、⑩题录、⑪索引、⑫产品样本及说明书。

三次文献：⑬手册、⑭综论、⑮教材讲义。

在分析123篇论文的引文中，我们可以看到没有⑧、⑩、⑪的引用出现，其余12种类型的文献则有程度不同地被引用。具体分布见图2所示。由图2可以看到，引文中期刊占第一位、图书占第二位，两者占总引文量的86%，其它十种类型仅占14%。应该指出，在上述十种类型引文所占的比重虽小，但其中产品样本及说明书、教材讲义、学位论文、内部资料等文献类型的被引用，这在其它基础研究论文中是较少见的，这正可说明我校是理工科性质大学的特征。

4. 引文的时间分布

我校学报78年复刊以来的123篇论文中，最早一篇的引文（期刊）是1912年发表的（是论文作者撰写专题综述而被引用的），最近的是当年（1982年初）发表的，前后时间跨度70年。若把引用文献发表的年度每10年区间作为一组，则可分为8组，然后进行统计分析，借以描述引文的时间分布状况，统计结果列于表5。

根据表5所列数据绘制成引文发表时间的分布图（见图3）。

从表5和图3可以看到大约有50%（引用率）的引文是在近10年来发表的，有78%左右

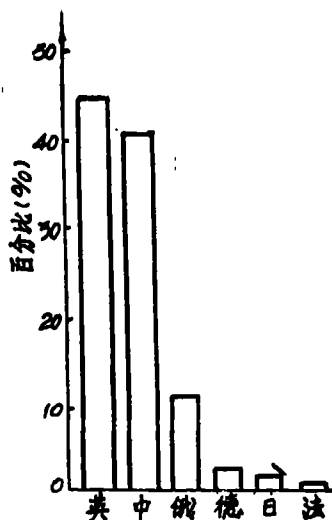


图1 引文语种的比较通信

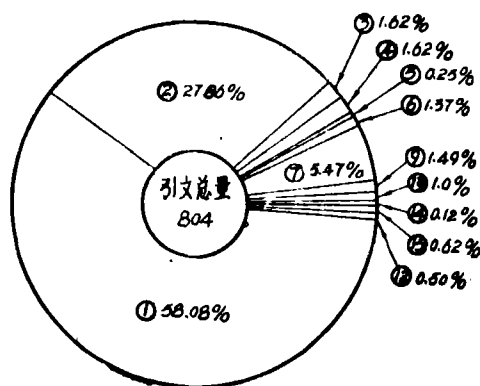


图2 引文类型的百分含量

表5 引文的时间分布

组 别	1	2	3	4	5	6	7	8	总
数量(篇)	80—82	70—79	60—69	50—59	40—49	30—39	20—29	10—19	计
年度 区间	80—82	70—79	60—69	50—59	40—49	30—39	20—29	10—19	
类 型	三月								
期 刊	25	210	120	69	19	8	15	1	467
图 书	10	86	76	39	10	2	1		224
专 利 说 明 书		8	2	1	2				13
科 技 报 告	2	9	1	1					13
学 位 论 文		2							2
学 术 会 议 论 文 集	3	7	1						11
内 部 资 料	3	38	3						44
文 摘		9	1		2				12
产 品 样 本 及 说 明 书		4							4
手 册		5	1	2					8
综 论		1							1
教 材 讲 义		5							5
总 计	43	384	205	112	33	10	16	1	804
%	5.35	47.76	25.50	13.93	4.11	1.24	1.99	0.12	(100%)

是近20年来发表的，有92%左右是解放后发表的。这就是说绝大部分的引文是解放后发表的。若以50年代这一年度区间发表的文献引用率为1，则60年代和70年代的引用率分别为1.83和3.43，这些数值近似地成为几何数列的增加。这种情况说明了科技文献随着时间的推移，被引用的新发表的文献大量地取代了旧的文献，这是文献老化的普通现象，它对于图书馆情报部门期刊等剔旧工作是具有一定的参考价值。还值得提出的，在我校学报78年复刊以来的四年中，被引用的科技报告、学位论文、学术会议论文集、内部资料、产品样本说明书和教材讲义等6种类型文献，有92%以上是70年代以后发表的，这说明我校师生和科技人员吸收现代新科学技术的情况和体现进行教学研究和应用研究的一个重要特征。

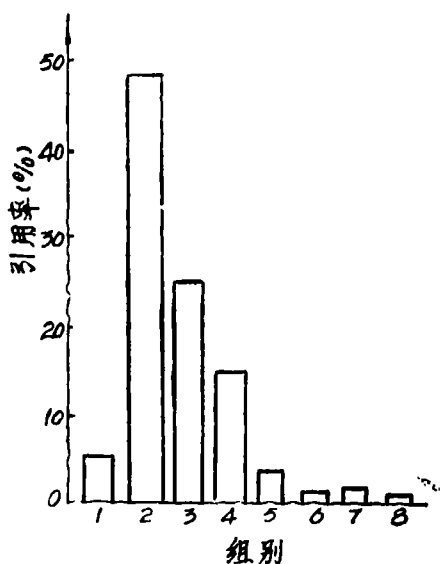


图3 引文发表时间的布分

结 论

1. 本文选择样本的方法及容量是较妥的。这可以通过如下验算来证实。

样本容量 W 可按下式计算:

$$W = \frac{V_1^2}{P_2} = \frac{(75.1)^2}{(6.7)^2} = \frac{5640}{45} = 125$$

计算结果说明要获得精确度 P 为6.7%，样本容量应定为125，本文样本容量为121，证明已具有较好的可靠性。

2. 我校学报自然科学论文以篇为单位，需要参考的文献数量范围是0—18篇，平均为6篇，引文的类型以期刊和图书占绝大多数，学位论文、科技报告，学术会议论文集和内部资料也占一定的比例。引文中出现了引用产品样本及说明书和教材讲义，体现了我校是理工科大学性质，教学与科研、理论与实践相结合的特征。

3. 引文的语种以英语为主，中文次之，说明大部分情报源来自英语期刊和图书。

4. 大约近20年内所发表的科学论文是主要的引用对象。

致谢：本文得到福建师范大学图书馆馆长金云铭教授的指导和帮助，谨致谢忱。

参 考 文 献

- [1] 福州大学学报，复刊词，(1)，(1978)。
- [2] 王崇德，图书情报工作，(2)，6(1982)。

CITATION ANALYSIS OF SCIENTIFIC PAPERS PUBLISHED IN FUZHOU UNIVERSITY JOURNAL

Liu Jiqing

Abstract

This article is written in an attempt to describe the utilization factors and characteristics of the literatures cited in the scientific papers published in Fuzhou University Journal after it resumed publication in 1978.

The preliminary conclusion is drawn from the citation analysis based on samples taken from the bibliographies of these papers. The numerical range of the literatures cited in a single one of the papers is from 0 to 18 and six on an average. 44.4% of all citations attributes to the literatures in English, constitutes a major portion. of the total citations 76.45% goes to the citations from literatures from foreign countries, and the rest to the country's own literatures.

The information sources, to a greater extend, are English books and journals, while a lesser part comes from lectures, degree-winning papers, text books, internal referenee materials, samples and informations of products, etc. All these show, on their own merits, the University's efforts on combining theory with practice, and teaching with research, which reflects the objective reality of the development of science and engineering.

Again, the faculty has an eye for the world's new development in science and engineering, so 78% of the citations are from the publications of the recent twenty years. It is proper to say that the statistical data are handled by strict hands, and the results of study of this paper are thus reliable.