Event No: 3242

American Society for Engineering Education

87th Annual Conference
Louisiana State University, June 25-28, 1979

Title: PITFALLS OF THE PITT STUDY

Author and Affiliation: HOMER I. BERNHARDT
UNIVERSITY OF PITTSBURGH

PITFALLS OF THE PITT STUDY

Much concern has been generated by a study done at the University of Pittsburgh which was sponsored in part by the National Science Foundation, with Professor Allen Kent as its principle investigator. The "Kent Study's" printed history is listed in the Bibliography. I will limit my discussion to the portions of the "Kent Study" (hereafter KS) that pertain to the Engineering Library, since it is under my purview, and further - the data collected for my internal study applies only to it. The Library Journal article which announced the initial findings of the KS was introduced by an editorial entitled "Pitt and the Pendulum" so it is only fitting that a paper disclosing its errors and faults be called "Pitfalls of the Pitt Study".

PITFALL 1: THE KS STATISTICS ARE BIASED

Webster's Third International Dictionary defines sampling as "the act, process, or technique of determining traits or attitudes of a whole population by collecting and analyzing data from a representative segment of it". Gregory Kimble in his book How to Use (and misuse) Statistics says "Samples that are not representative of the population of interest are said to be biased." He suggests "randomness" as the most straightforward way to avoid bias in a sample..." A random sample presents the opportunity for each item or combination of items to

have the equal chance of being selected. As the size of the sample (provided it is random) approaches the size of the population, it becomes a better estimate of the whole. On page 53 of the KS's Final Report they point out that the libraries were "pre-sampled" for 1 week, at 1 hour intervals, in order to determine the distribution of usage. The resulting periods of "heavier hours of usage", i.e., those times with more persons in the library, determined the observation periods. Not only does this lead to a biased sample, but their assumption that more people means more journal usage is just not true. When our library is crowded - it is crowded, and there are no tables or study carrels available for serious or intensive use of journals. This explains both the quantitative and qualitative differences in the journals used during the day, from those used at night and on weekends.*

During the day reshelvings consist of many current journals and a few bound volumes, while the opposite occurs at night and on weekends. This indicates to me that persons browse or hunt down specific citations during the day, and use the evenings and weekends for in-depth journal utilization.

Additional proof that large numbers of persons have no bearing on journal usage can be found in Appendix 1. Here attendance figures are compared to unique title uses. The months with the poorest attendance show the greatest percentage of

*Data from reshelvings (no observations) has been gathered and tabulated.

unique title uses per patron.

Another pitfall of the KS is their assumption that journal usage is static and constant, the basis of the sample in the Final Report which I already mentioned. Our own in-house study of tabulations from reshelvings provides some interesting information. Comparing unique titles used (first time uses of journals) for the first 6 months of 1977/78 with those of 1978/79, we found that although 308 titles were common to both time frames, there were 74 titles used in the first period but not the second; and there were 120 titles used in the second period but not the first - two similiar periods of time and a variance of 194 titles. Not only does this destroy their sampling methodology and their biased sample, but it points out the peculiar and unique way in which journals are used, and negates the purpose of their journal study:

> "a. to develop a methodology that would provide librarians with a relatively simple mechanism for discerning patterns of usage; and
>
> b. to test this methodology by studying the patterns of usage in the Science and Engineering libraries on the University of Pittsburgh campus."*

Journals are not bestselling novels or classic reference books whose use is fairly constant and largely predictable. As a rule, a user seeking a particular journal article will not accept a substitute. Another article on the same

*Final Report, p. 51

subject will not satisfy or take the place of the one which the patron wants, as a text on FORTRAN might for a person needing some basic information on the "do-loop".

A sample must be representative and true of the base from which it is drawn if the conclusions from its analysis are to be meaningful. It follows that if the base is poor, the sample will be poor. This is the case in respect to the Engineering Library's journal holdings vis-a-vis what the KS calls "usage by age". The collection has an average age of 6.1 years, a median age of 10.95 years, and a mean age of 14.4 years, which is hardly a good base to sample to study the "aging of engineering journals". This, coupled with the faulty sampling techniques previously pointed out casts grave doubts on the validity of the results reported in the Library Journal by Galvin and Kent, and in the Final Report to NSF.

PITFALL 2: THE HOLDINGS ARE REPORTED INACCURATELY

Perhaps the most precarious pitfall of the KS lies in the figures, and the manner in which they were presented. Appendix 2 gives a complete picture of the Engineering Library's journals by category. The KS's revised Progress Report says "The collection was taken to be 'current subscriptions,' so that the figures show a higher usage than would be true if we took the collection to be 'all titles in the collection.' " (p.23) The percent of use, i.e., unique title uses, is 8.4. The LJ article agrees with this percentage. But the 8.4% is the

"sample" percentage, not the extrapolated one. In the Final Report the number of "titles in the collection" jumps from 687 to 1,643, and the percent of use drops to 6.8. Once again the 6.8% is the unextrapolated version which rises to 84% (p.63) when it is calculated for titles that will be used 2 or more times a year.

What has happened? Simply, the rules were changed in the middle of the ballgame. The Final Report says that the "titles in the collection" are now taken to be the "official figures" (p. 59). Thus, they include microforms, standing and continuation orders (both of which are cataloged and kept with the book collection), gifts, exchanges, I & A's, and material used by the library staff, e.g., BIP, PTLA, CBI, etc., items that were not sampled. Current subscriptions somehow rose from 687 to 737 (p. 66).

The University Libraries' serial record tapes were used to generate the "official figures". The tapes were designed to produce a serial record - not statistical data. This is why material that was excluded from the "sample" was included in the "official figures", but the erroneous data doesn't end with this. Entries can have more than one holdings line as is shown in Appendix 3. Thus, there is a difference between "total holdings" and "total titles". A title can also have more than one entry when it is received both in hardcopy and microform in lieu of binding (Appendix 4). Consequently, there is also a difference between "total titles" and actual

"total titles". A serials printout for the Engineering Library, dated 11/15/78, lists it with 1,678 "total holdings" and 1,529 "total titles" - already a difference of 149. A detailed breakdown, which was done manually, is shown in Appendix 5. To summarize it: the printout contains 149 duplicate holdings lines; 355 titles which were excluded from the sample; 214 journals with cosmetic title changes (The change in title did not create a substantive change in the content of the journal, e.g., IRE to IEEE, Catalysis Reviews to Catalysis Reviews: Science and Engineering, etc., nor did the volume numbering change); and 25 items which were bad data. This totals up to 743 which can immediately be deducted from 1,678 leaving a balance of 935. If the 158 gift/exchange items are deleted, we are left with a count of 777 viable journals, which includes 84 ceased or cancelled titles, and 40 with real title changes. If we discount these dead runs (the 84 and 40), we are left with our current subscriptions - 653 (actually 633 since it includes 10 titles that we did not own at the time of the KS, and 10 suspended journals). A lot depends on what you want to count, but there is no hesitation on my part to exclude titles that were not sampled or those with multiple holdings and entries. If we are to be penalized for making gifts (house organs, newsletters, etc.) available, maybe we should throw them out instead.

Why did the "Kent Study Team" knowingly use incorrect data when they knew it was erroneous midway through the study?

The explanation given in the Final Report (pp. 59-60) is not satisfactory:

> "There was considerable discussion of the precise number of titles in the journal collections for the science and engineering libraries. The figures from the official serials record are listed under 'Official Figures' above. The departmental librarians said that there were fewer titles actually in the collections, and their figures are listed under 'Librarian's Figures'. In some cases, the discrepancy is not insignificant, as in the Engineering Library... However, in the interests of consistency, the decision was made to use the official list wherever possible. ... We have not recalculated our figures, which are based on the 'official' count. We have included the librarians' figures so that the interested reader may be in a position to recalculate the percent of usage if he wishes to do so. ..."

How can they justify that statement when the "official" serial record count was explained to them at the time they received it?

The draft of the Final Report has one of the professors using 1,485 for the "number of titles" in the Engineering Library, and another one using 1,643 (pp. 108 & 57). The "subscription costs" include $ 9,072 for indexes and abstracts which were not part of the study (p. 108). When both of these discrepancies were pointed out to them, the figure for the "number of titles" was changed from 1,485 to 1,643 (Final Report, p. 107), but the amount for "subscription costs" was only footnoted! Why is it that that the higher cost figures are highlighted in the body of the text or on the tables, and modified to lower figures by footnotes? The "subscription cost/use" for the Engineering Library is shown to be $6.81 on pages

67 and 107 of the Final Report. Footnote 7 on page 111 drops it to $5.43. The "average cost per use" is given as $9.04 on pages 67 and 111, but is lowered to $7.21 in footnote 7 on page 111. Although these figures are just as wrong as the result of their faulty sampling and methodology, it is important to bring them to your attention in this light because of the widespread and national publicity that the KS has been given.

Professor Jasper G. Schad in the May, 1979, issue of the Journal of Academic Librarianship states:

> "Despite its impressive array of statistical data and intricate formulas, the study is based on incorrect assumptions and incomplete data that lead to meaningless conclusions.
>
> Simply stated, the Pittsburgh study does not demonstrate comprehension of the purpose of an academic research or university library."

Professor Melvin J. Voigt in the same issue of the JAL stated in letters to Professors Kent and Galvin:

> "The pattern of scientific journal use that your study reportsis so different from what one normally finds in scientific libraries that it must be assumed that (1) Pittsburgh is different, with journals being used in other libraries in the area or in departmental or office collections, or (2) there is little research activity in these fields, or (3) that the methodology is faulty. ...
>
> In summary, your first study does not cover the scholarly use of research materials in a university research library, admittedly a very difficult task and to my knowledge, one not attempted thus far. The second study, for reasons impossible to fully identify from the report, presents findings that are contrary, not only to my experience, but I am certain, to that of most librarians involved with active scientific and technical libraries. I must, therefore, repeat that this type of report can and will do unmeasurable harm to the scholarly research supported so precariously by research libraries today.

I am taking the liberty of sending this letter and my earlier one to a number of persons who, I believe, will be interested in my views."

The faculty library representatives at the University of Pittsburgh also criticized the KS in a report that was approved at their Winter meeting. Pitt's Senate Library Committee also has the KS under investigation. The results will soon be released. Interested parties should contact this speaker for more information.

SUMMARY

How can we measure journal usage? Should we only use numbers and ignore qualitative factors? What is important? Do 50 uses of a popular title (Aviation Week) outweigh 10 uses of a scholarly one (Journal of Fluid Mechanics)? How vital is it to our patrons to have the ability to browse? What about the frequency of publication - how can we equate the number of uses of a weekly with a quarterly, a monthly, etc.? How do you compare journals with many articles with those that only have a few? Which journal is more important - one that is used 100 times two months out of the year, or one that is used 2-3 times every month? Does the time delay to users when photocopies have to be ordered come into play? What about their cost - if paid by the patron, their department, or the library? These are just a few of the questions that we must answer.

I agree that there is a "core collection of journals". Many persons have used Bradford's Distribution and citation analysis to examine this issue in depth, so its announcement in

the KS in nothing new. The fundamental concept of a research library is the development of its collections beyond the realm of this core, in order for it to support the educational and research needs of its patrons and the unit it was designed to serve. Once this concept is diluted, its ability to function at the proper academic level will cease.

In order to gather some raw data for our internal study, our staff fills out a simple little form (Appendix 6) for each journal to be reshelved. The number of uses are recorded on a "monthly data sheet" (Appendix 6).* Unique titles uses are taken from this and placed on a "12 month composite" (Appendix 7).* This gives us an annual representation of the titles used in that particular academic year. This data sheet is adaptable to a "12 year composite" (Appendix 7) by changing its headings. Both the 12 month and 12 year composites discount number of uses, since they address unique uses. We will be able to see some obvious trends, but again it is only a quantitative measure. The development of journal collections are slow and continuing processes. There is no quick and simple mechanism to use as the KS suggests.

In conclusion the KS's sample was biased, and the assumptions on which is was based were not correct. Holdings figures were inaccurately reported. Thus, their results and conclusions do not represent the actual. Usage is underestimated, and cost/use overestimated. Aging is not a true representation. There is no "cost-benefit" model to follow - the title of their study.

*Generated from our own computerized file.

APPENDIX 1

DATE	ATTENDANCE*	UNIQUE** TITLES USED	UNIQUE TITLES USED PER 1,000 PATRONS
SEP'77	17,468	113	6.5
OCT	23,430	143	6.1
NOV	22,616	179	7.9
DEC	15,157	121	8.0
JAN'78	15,280	108	7.1
FEB	21,913	140	6.4
MAR	27,363	204	7.5
APR	22,474	209	9.3
MAY	7,188	245	34.1
JUN	6,960	204	29.3
JUL	6,171	135	21.9
AUG	4,743	179	37.7
SEP	18,449	170	9.2
OCT	26,769	214	8.0
NOV	25,449	207	8.1
DEC	15,780	148	9.4
JAN'79	22,316	177	7.9
FEB	24,284	176	7.3

*Taken from "Tattle-Tape" counter.

**Taken from "in-house study".

APPENDIX 2

REFERENCE ITEM	TITLES IN THE COLLECTION	CURRENT SUBSCRIPTIONS	SAMPLE USAGES	PERCENT OF COLLECTION
PR-2 **	687 (p.23)	687 (p.23)	58*	8.4 (p.24)
LJ **	687 (p.2320)	687 (p.2320)	58 (p.2320)	8.4 (p.2320)
FR **	1,643 (p.59)	737 (p.66)	172 (p.57)	6.8 (p.60)

*687 X 8.4 = 57.708

**Initials refer to items in the bibliography.

APPENDIX 3

TITLES WITH MORE THAN ONE HOLDINGS LINE

AMERICAN INSTITUTE OF MINING ENGINEERS
-TRANSACTIONS
V1,1871-V59,1918. CONTINUED AS AMERICAN INSTITUTE OF MINING AND METALLURGICAL ENGINEERS. TRANSACTIONS.
*
UE V1-59.
UE INDEX V1-35(1871-1905),V36-55(1905-15).
PITT NO. 0438000005

AMERICAN INSTITUTE OF MINING, METALLURGICAL AND PETROLEUM ENGINEERS
-TRANSACTIONS
V205,1956- FORMERLY AMERICAN INSTITUTE OF MINING AND METALLURGICAL ENGINEERS. TRANSACTIONS.
*
UE V205-
UE INDEX V118-243(1936-68)
PITT NO. 0438300009

AMERICAN FOUNDRYMEN'S ASSOCIATION
-TRANSACTIONS
V12,1904-V55,1948. CONTINUED AS AMERICAN FOUNDRYMEN'S SOCIETY. TRANSACTIONS. FORMERLY AMERICAN FOUNDRYMEN'S ASSOCIATION. JOURNAL.
*
UE V42-
UE INDEX V38-48(1930-40)
PITT NO. 0419400001

AMERICAN FOUNDRYMEN'S SOCIETY
-TRANSACTIONS
V56,1948- FORMERLY AMERICAN FOUNDRYMEN'S ASSOCIATION. TRANSACTIONS.
*
UE V56-57,59-
UE INDEX V49-58(1941-50),V59-68(1951-60)
PITT NO. 0419700005

APPENDIX 4

TITLES WITH MORE THAN ONE ENTRY

PRODUCTION ENGINEER (INSTITUTON OF PRODUCTION ENGINEERS)
LONDON
V1,1921/22- 1921-25 AS INSTITUTIONS PROCEEDINGS. 1926-59 AS INSTITUTION'S JOURNAL.
*
HI V39INC.
UE PAPER COPY KEPT UNTIL RECEIPT OF MICROFILM.
PITT NO. 6611100008

THE PRODUCTION ENGINEER (MICROFILM) (INSTITUTION OF PRODUCTION ENGINEERS)
V39,1960- MONTHLY. V1,1921/22-V38,1959 AS THE PRODUCTION ENGINEER JOURNAL. SUPPLEMENTS ACCOMPANY SOME NUMBERS.
UE MICROFILM V44-
PITT NO. 6611100016

AVIATION WEEK AND SPACE TECHNOLOGY
NEW YORK, NY
V72NO2,JA1960- WEEKLY. FORMERLY AVIATION WEEK, INCLUDING SPACE TECHNOLOGY.
*
HI V72-73,92-93INC.
UB V71-91.
UP CURRENT YR. ONLY
UE PAPERCOPY KEPT UNTIL RECEIPT OF MICROFILM
PITT NO. 1055400001

AVIATION WEEK AND SPACE TECHNOLOGY (MICROFILM)
NEW YORK, NY
V72NO2,JA1960- WEEKLY. FORMERLY AVIATION WEEK, INCLUDING SPACE TECHNOLOGY.
UE MICROFILM V72-
PITT NO. 1055400010

APPENDIX 5

ANALYSIS OF THE ENGINEERING LIBRARY'S PRINT-OUT DATED 11/15/78

TYPE OF ENTRY	NUMBER OF ENTRIES
CONTINUING PAID	
CURRENT SUBSCRIPTIONS	646*
CEASED/CANCELLED	84
TITLE CHANGE (COSMETIC)	203
TITLE CHANGE (REAL)	40
CONTINUING GIFT/EXCHANGE	
CURRENT SUBSCRIPTIONS	154#
CEASED/CANCELLED	8
TITLE CHANGE (COSMETIC)	11
TITLE CHANGE (REAL)	3
OTHER	
INDEXES & ABSTRACTS	121
MICROFORM	75
C.O./S.O.	150
INTERNAL USAGE	9
ENTRIES NOT MARKED	
DO NOT OWN OR HOLD	21
DUPLICATE ENTRIES	3
CROSS REFERENCE	1

*Includes 3 suspended titles, and 10 new ones not yet received at time of the Kent Study.

#Includes 7 suspended titles.

APPENDIX 6

RECORDING FORM

TITLE _____

_____

VOL _____ YEAR _____ PART _____

F _____ G _____ U _____ OTHER _____

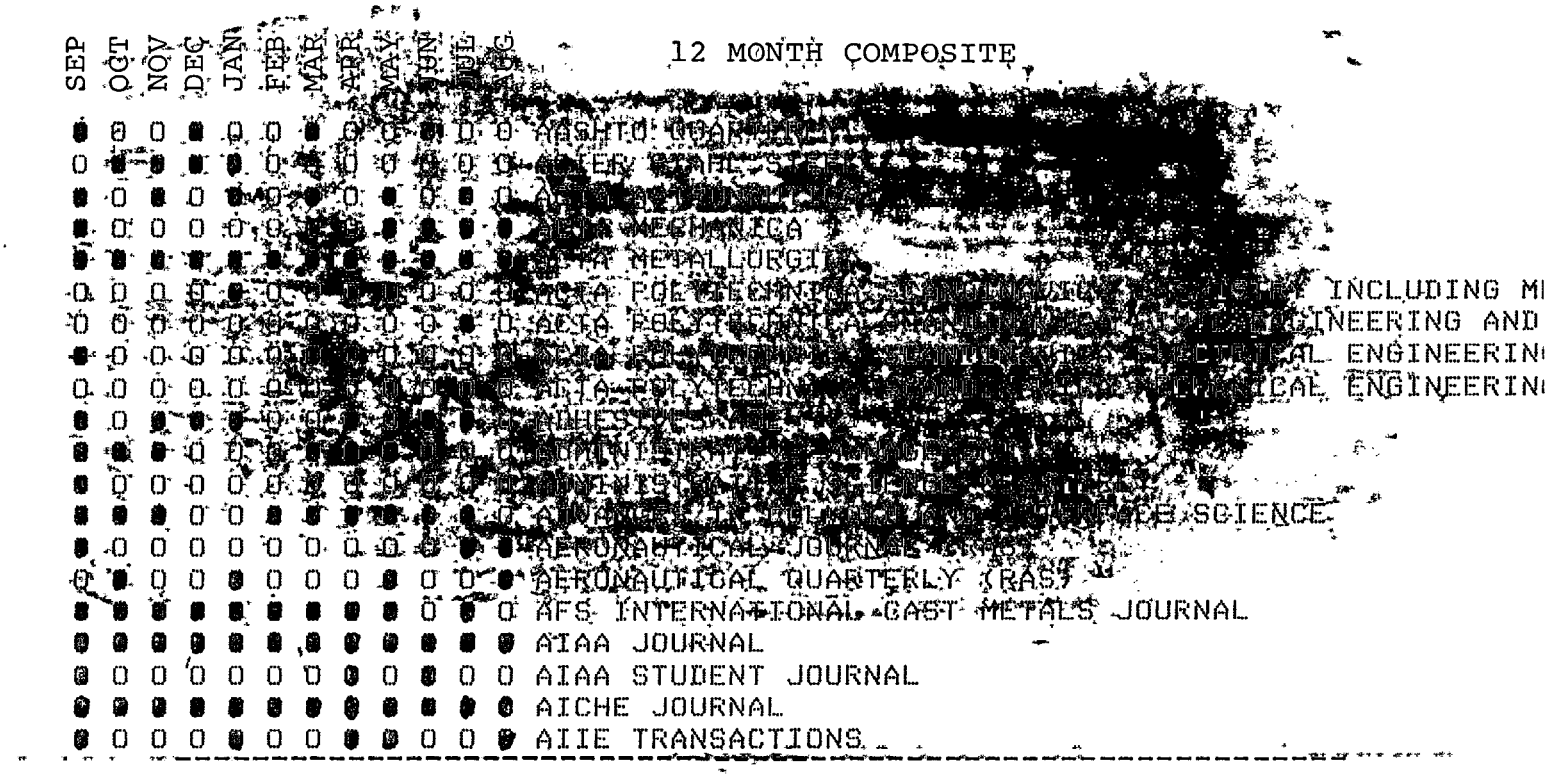
OUT:

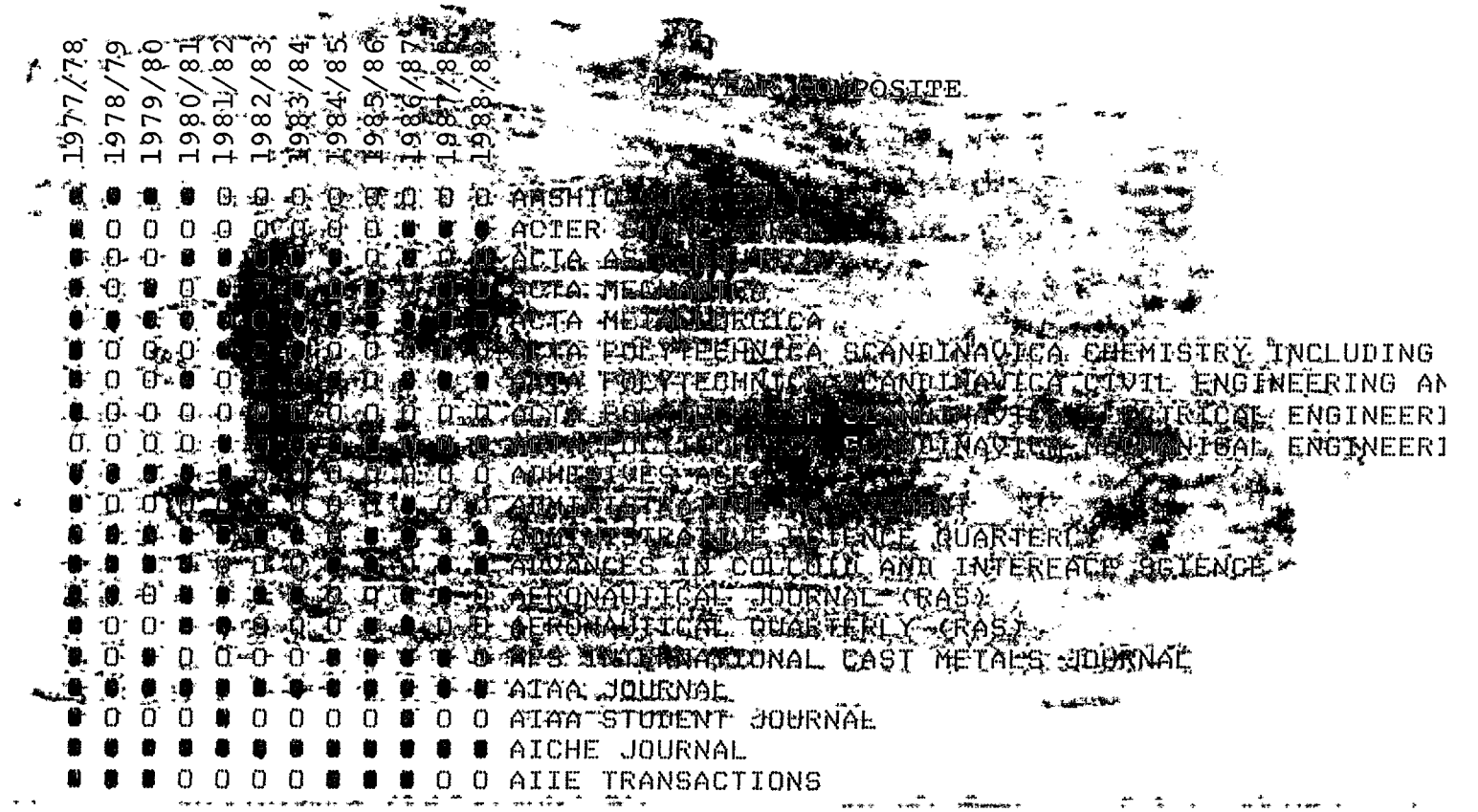
IN:

MONTHLY DATA SHEET

AASHTO QUARTERLY
ACIER STAHL STEEL
ACTA ASTRONAUTICA
ACTA MECHANICA
ACTA METALLURGICA
ACTA POLYTECHNICA SCANDINAVICA METALLURGY SERIES
ACTA POLYTECHNICA SCANDINAVICA BUILDING CONSTRUCTION
ACTA POLYTECHNICA SCANDINAVICA SERIES
ACTA POLYTECHNICA SCANDINAVICA SERIES
ADHESIVES
ADMINISTRATIVE
ADMINISTRATIVE
ADVANCES IN
AERONAUTICAL
AERONAUTICAL
AFS INTERNATIONAL CAST METALS JOURNAL
AIAA JOURNAL
AIAA STUDENT JOURNAL
AICHE JOURNAL
AIIE TRANSACTIONS

APPENDIX 7


SEP OCT NOV DEC JAN FEB MAR APR MAY JUN JUL AUG
12 MONTH COMPOSITE
AASHTO QUARTERLY
ACTA MECHANICA
ACTA METALLURGICA
ACTA POLYTECHNICA SCANDINAVICA CHEMISTRY INCLUDING
ACTA POLYTECHNICA SCANDINAVICA CIVIL ENGINEERING AND
ACTA POLYTECHNICA SCANDINAVICA ELECTRICAL ENGINEERIN
ACTA POLYTECHNICA SCANDINAVICA MECHANICAL ENGINEERIN
ADHESIVES AGE
ADMINISTRATIVE SCIENCE QUARTERLY
ADVANCES IN COLLOID AND INTERFACE SCIENCE
AERONAUTICAL JOURNAL (RAS)
AERONAUTICAL QUARTERLY (RAS)
AFS INTERNATIONAL CAST METALS JOURNAL
AIAA JOURNAL
AIAA STUDENT JOURNAL
AICHE JOURNAL
AIIE TRANSACTIONS


1977/78 1978/79 1979/80 1980/81 1981/82 1982/83 1983/84 1984/85 1985/86 1986/87 1987/88 1988/89
12 YEAR COMPOSITE
AASHTO QUARTERLY
ACTER
ACTA MECHANICA
ACTA METALLURGICA
ACTA POLYTECHNICA SCANDINAVICA CHEMISTRY INCLUDING
ACTA POLYTECHNICA SCANDINAVICA CIVIL ENGINEERING AN
ACTA POLYTECHNICA SCANDINAVICA ELECTRICAL ENGINEERI
ACTA POLYTECHNICA SCANDINAVICA MECHANICAL ENGINEERI
ADHESIVES AGE
ADMINISTRATIVE SCIENCE QUARTERLY
ADVANCES IN COLLOID AND INTERFACE SCIENCE
AERONAUTICAL JOURNAL (RAS)
AERONAUTICAL QUARTERLY (RAS)
AFS INTERNATIONAL CAST METALS JOURNAL
AIAA JOURNAL
AIAA STUDENT JOURNAL
AICHE JOURNAL
AIIE TRANSACTIONS

APPENDIX 8

	1977/78 UNIQUE TITLES USED*						1978/79 UNIQUE TITLES USED*					
	THIS MONTH ALONE		ADDED THIS MONTH		CUMULATIVE FOR YEAR**		THIS MONTH ALONE		ADDED THIS MONTH		CUMULATIVE FOR YEAR**	
	#	%	#	%	#	%	#	%	#	%	#	%
SEP	113	18%	113	18%	113	18%	170	27%	170	27%	170	27%
OCT	143	23%	99	16%	212	33%	214	34%	121	19%	291	46%
NOV	179	28%	82	13%	294	46%	207	33%	60	10%	351	55%
DEC	121	19%	28	4%	322	51%	148	23%	25	4%	376	59%
JAN	108	17%	28	4%	350	55%	177	28%	28	4%	404	64%
FEB	140	22%	32	5%	382	60%	176	28%	24	4%	428	68%
MAR	204	32%	23	4%	405	64%						
APR	209	33%	33	5%	438	69%						
MAY	245	39%	37	6%	475	75%						
JUN	204	32%	17	3%	492	78%						
JUL	135	21%	7	1%	499	79%						
AUG	179	28%	12	2%	511	81%						

*On a base of 634 paid subscriptions.

**In the first 6 months of 1977/78 and 1978/79 the following observation is noted: 74 titles used in 1977/78 have not yet been used in 1978/79; 120 titles used in 1978/79 were not used in 1977/78; 308 titles have been used in the first 6 months of both years.

BIBLIOGRAPHY

(PR-1) Kent, A., et al., A Cost-Benefit Model of Some Critical Library Operations in Terms of Use of Materials, Progress Report, April 1, 1977, 42p.

(PR-2) Kent, A., et al., A Cost-Benefit Model of Some Critical Library Operations in Terms of Use of Materials, Progress Report, April 1, 1977 (Revised April 29, 1977), 42p.

(LJD) Galvin, T.J. and Kent, A., "Use and Non-Use of a University's Collections: A Progress Report on the Pittsburgh Study", July, 1977 (Scheduled for Publication - Library Journal - November 15, 1977), 12p.

(LJ) Galvin, T.J. and Kent, A., "Use of a University Library Collection: A Progress Report on a Pittsburgh Study", Library Journal, November 15, 1977, pp. 2317-20.

(FRD) Kent, A., et al., A Cost-Benefit Model of Some Critical Library Operations in Terms of Use of Materials, Final Report (Draft - December, 1977).

(FR) Kent, A., et al., A Cost-Benefit Model of Some Critical Library Operations in Terms of Use of Materials, Final Report, April 15, 1978, 243p.