

Legibility of Library of Congress Cards and Their Reproductions

REPRODUCTION in book form of the *Library of Congress Catalog of Printed Cards* by the photolithoprint process, a project nearing completion by the Association of Research Libraries with the technical aid of Edwards Brothers, is an achievement which undoubtedly advances solution of some of the serious problems involved in library purchase, maintenance, and use of this important tool. By October 1945 it is expected that all subscribing libraries will have received the complete set of 176 volumes of this work, containing reproductions of approximately two million Library of Congress cards. Library demand for these volumes already exceeds available supply, and Edwards Brothers plans to reprint the entire work after the war if enough subscriptions are obtained to support a second edition. A supplement to the present edition, to be published immediately after the war, is definitely projected.¹

To assure proper distribution of the reproduced printed catalog outside the United States, the Rockefeller Foundation has contributed \$37,500, through a grant to the American Library Association, toward the cost of producing this work. In describing the project, the foundation explained that "expenses for supplying an entire set of Library of Congress cards formerly averaged about \$7000. Present procedure, which leaves data from the cards still entirely legible and available in familiar book

form, has cut this cost to \$750 or less per set." The Rockefeller grant thus will make available to fifty selected institutions abroad "this most important single bibliographical resource developed by American librarianship."²

Librarians' Reactions to Reproductions

Following preliminary approval and subscription to the new book form of the Library of Congress catalog, adverse criticism began to be voiced as successive volumes rolled off the press, both by library administrators and by catalogers. Objections were directed chiefly at the illegibility of some of the reproduced material as compared with the original cards. A survey conducted by Jens Nyholm, then of the University of California Library³ and now of Northwestern University Library, showed that a majority of the twenty-three respondents, for the most part directors of large university and public libraries, concurred in the judgment that the reproductions were unsatisfactory with respect to legibility. The chief sources of reading difficulty attributed to the reproduced cards were excessive reduction in type size, blurring, and reduction of contrast in the offset material. As a result of these factors, it was asserted, more time was required for reading the small print, errors in ordering cards from serial numbers were more likely to

² *Rockefeller Foundation, Annual Report. 1942, p. 219-20.*

³ "Summary of Comments on a Catalog of Books Represented by Library of Congress Printed Cards in Response to an Inquiry by the University of California Library." Distributed in mimeographed form under date of June 5, 1943.

¹ "Edwards Brothers Expands Its Field of Scholarly Publishing." *Publishers' Weekly* 146:314-18, July 29, 1944.

occur, and eyestrain was more noticeable in using the reproductions than when working with the original printed cards.

In opposition to this judgment, several catalogers interrogated in the course of the survey reported that they had experienced no difficulty in reading the entries and no eyestrain even after several hours of continuous work. In assessing the quality of the reproductions after some fifty volumes had been published, Bishop⁴ likewise noted that some persons who had worked steadily with the reduced material had reported no eyestrain. While conceding that the reproduction of the printed portion of the cards is not entirely satisfactory in quality and that "the reduction is perhaps too great for complete legibility in all cases," Bishop concluded that "for all ordinary purposes of consultation the book is perfectly legible . . . it is a perfectly good working tool." As spokesman for the committee which carried out the project, he explained that a lesser reduction would have increased the cost of the entire work by at least \$200, or almost a third, and he stressed the savings to the profession to be effected by the reasonable purchase price of the new tool.

Nyholm, on the other hand, indicated that the considerable savings anticipated through use of the bound volumes by libraries maintaining depository sets of Library of Congress cards had not been realized. As evidence he cited the fact that only six libraries in a group of sixty-one had decided, after acquiring some volumes of the bound reproductions, that they would replace their depository sets with the new tool and return the former to the Library of Congress. Nyholm implied that a majority of the fifty-five libraries reporting to the Library of Congress that they would not return their depository sets, or that they were undecided in the matter, were

deterred from taking this action by the unsatisfactory quality of the reproductions of the cards. Some support for this assumption may be found in the responses relating to this question which were obtained from the smaller sample of twenty-three librarians who participated in the Nyholm survey.

As pointed out by Tate⁵ in discussing various mechanical techniques designed for coping with the problem of the rapidly accelerating growth of research libraries, the reaction of the ultimate consumer is a factor vital to successful adoption and use of any new tool. He refers to library experience with miniature photographic facsimiles: "It has been found that some users eagerly adopt them while others are equally vociferous in rejecting them in any form." One source of uncertainty in attempting to gauge consumer reaction in advance of use inheres in the fact, observed by Tate, that "there is a vast difference between a laboratory sample and the manufactured product." He concludes that consumer acceptance cannot be determined "arbitrarily through discussion and speculation; a conclusive answer can only be obtained through a test project."

As far as can be ascertained, no test project was conducted to determine the relative legibility of the L.C. cards and their reproductions before the volumes were delivered to subscribers. In preliminary estimates of the probable savings to be effected by libraries through purchase and use of the reproduced Library of Congress catalog, an approximate equivalence in legibility of the printed cards and their reproductions, as far as this factor may affect working efficiency, seems to have been assumed. Consumer reactions, after examination and use of the published bound volumes, cast doubt upon

⁴ Bishop, William Warner. "Notes on the Library of Congress Catalog of Printed Cards." *Library Journal* 68:869-71, Nov. 1, 1943.

⁵ "Controversial." A review by Vernon D. Tate of Fremont Rider's book, *The Scholar and the Future of the Research Library*. *Library Journal* 69:1046, Dec. 1, 1944.

this assumption. Unwilling to base decision regarding the advisability of replacing a set of the printed L.C. cards with the bound reproductions upon users' opinions alone, especially since there was conflict among them, Carl M. White, director of the Columbia University Libraries, requested the writer to obtain some objective data on this problem.

Experts' Judgments on Legibility

Nyholm, in the mimeographed report of his survey referred to above, included a statement by Matthew Luckiesh, director of the Lighting Research Laboratory of the General Electric Company, on the visibility of the type used in the book catalog. Regarding the reduction of the full-size 10-point type used on the L.C. cards, Luckiesh states: "The reproductions of the original 10-point type result in equivalent type-size from 5.4 to 6.5. This is extremely small type to be read for long periods, particularly when there are various details to be distinguished." In commenting on the reduction of the full-size 8-point type also used on L.C. cards, Luckiesh reports: "The equivalent type-size of the reproductions varies from 4.5 to 5.7 as determined by measurements of relative visibility. . . . It is little short of outrageous to ask anyone to pore over type of this size. In addition to these small resulting sizes, we have the blurring, filling in, and loss of detail in various reproductions." He concludes that "even if the reproduction is perfect in each case, the smallness of the equivalent type-size places an extreme burden upon eyes and upon the entire human seeing-machine."

To obtain further expert opinion, the writer consulted Miles A. Tinker, of the University of Minnesota. In a study concerned with preservation of newspaper accounts of historical events, Tinker and Paterson⁶ found that reduction to 50 per

⁶ Tinker, Miles A., and Paterson, Donald G.

cent of the original 7-point text lessened the readability significantly, although the text could still be read by college students. Just beyond a size reduction of 50 per cent, the legibility curve dropped sharply. With a reduction to 30 per cent of original size, the type was blurred and could be read only with great difficulty. Later studies by the same experimenters⁷ showed that speed of reading is retarded when type sizes less than 9-point are used. In experimenting with eye movements involved in reading, these authors found⁸ that 6-point type when read by college students yielded a much less efficient oculomotor pattern than 10-point type. Photographs of eye movements taken while reading was in process showed that with the 6-point type the number of fixations was increased, the span of perception was decreased, the time per fixation was increased, total perception time was greatly increased, and the frequency of regressive movements was slightly increased.

When asked to give his opinion on the legibility of the reproductions of the L.C. cards, Dr. Tinker replied that he believed his experience in the field warranted judgment without a special experiment. Upon examination of material in the bound volumes, his findings⁹ were:

(1) The reduction in the size of print in the offset material produces a poorly readable copy. (2) Add to this the blurring and reduction of contrast which occurs in this offset printing and we have a copy that not only requires excessive visual effort to read but also a copy that will be read less efficiently from the viewpoint of reading time and possible errors. Add to this the subjective feeling of discomfort and irritation on the part of the reader. . . . With the best of eyes this mate-

"Studies of Typographical Factors Influencing Speed of Reading." *Journal of Applied Psychology* 16:525-31, October 1932.

⁷ Paterson, Donald G., and Tinker, Miles A. *How to Make Type Readable*. New York, Harper and Brothers, 1940. 209p. (Size of Type, p. 29-37.)

⁸ Paterson, Donald G., and Tinker, Miles A. "Influence of Size of Type on Eye Movements." *Journal of Applied Psychology* 26:227-30, April 1942.

⁹ Reported in a letter to the writer dated Nov. 2, 1943.

rial will cause eye fatigue and lessened efficiency by the end of one hour.

In contrast to these judgments was that of Leonard Carmichael, of Tufts College, and Walter F. Dearborn, of the Psycho-Educational Clinic of Harvard University. During recent years the Tufts Laboratory has been actively engaged, in cooperation with Dr. Dearborn, in the study of fatigue in connection with reading. The technique used in these visual fatigue studies involves electrical recording of all eye movements during prolonged reading periods, supplemented by measures of comprehension and other relevant variables. These studies seem to indicate, report Carmichael and Dearborn,¹⁰ that the visual mechanism is amazingly resistant to fatigue in reading. (They cite the finding, for example, that six hours of continuous reading of microfilm showed no measurable decline in visual efficiency or "fatigue.") The feeling tone accompanying reading, they point out, may change when the visual mechanism is apparently not being fatigued; that is, its effective operation is not impaired. Thus, readers will often assume that their visual mechanism is being harmed because they are provoked by or esthetically unsatisfied by a particular type of format.

Regarding the legibility of the reproductions of the L.C. cards, Carmichael and Dearborn state:

Frankly, we do not believe that anyone will be willing to make a final statement concerning the legibility or fatiguing nature of any special form of reproduced printed material without specific study, particularly in the case of material . . . in which type size, blurring of contours, and decrease contrast are all factors. . . . It is our guess that material of the sort on the page which you enclose could be read for hours by a normal individual, providing reading light is satisfactory, without visual harm. We are equally certain that if this reading were done by an unselected group

¹⁰ Letters to the writer under date of Oct. 28, 1943, and Apr. 19, 1945.

of subjects a certain number of them would be willing to report headaches, soreness in the lids, and so forth. Suggestion plays a large part in our estimate of reading ease.

In commenting on these psychological factors involved in reading efficiency, Edwin G. Boring, director of the Psychological Laboratory at Harvard University, agreed¹¹ that the eyes will stand a great deal without impairment of function but notes that "when people have to make constant adjustment to get their eyes to work there may be an emotional strain which does not show in error. That complicates the problem." Calling attention to the change in the eyes that occurs in the average person around the age of forty, Boring points out that there must be a great many catalogers over forty-five years who would use the L.C. card reproductions and that findings obtained with young eyes may be misleading.¹² The inclusion of foreign language material on the L.C. cards, according to Boring, is another variable that should be taken into account when comparing the cards with their reproductions. This expert recommended that a special study be undertaken to determine to what extent the use of the reproductions, as compared with the original L.C. printed cards, would affect the working efficiency of catalogers.

The opinions of these four specialists, while not unanimous, gave some support to the librarians who had rated the reproductions as unsatisfactory with respect to legibility. On the question of visual fatigue, opinion was divided. With regard to working efficiency, while one expert stated that the reproduced material would be read less

¹¹ Letter to the writer, Oct. 27, 1943.

¹² A recent study by Bryan and Curtiss showed that the median age of 89 catalogers in eleven large public libraries and 125 catalogers in eight large university libraries was approximately forty years. About four-fifths of each group were between the ages of twenty-five and fifty-four, with approximately a third between thirty-five and forty-four; about a seventh of each group were over fifty-four years of age. ("Report to Cooperating Libraries on Results of Columbia Cataloging Examination Project." School of Library Service, Columbia University, 1943. Mimeo. 14p.)

efficiently in terms of reading time and possible errors, no estimate was given as to the probable extent of the decreased efficiency. From the standpoint of library economy, this is an important consideration, since an increase in reading time and errors on the part of professional catalogers would increase the use cost of the bound reproductions and might cancel or even outweigh the higher cost of purchase, filing, and storage of the printed cards.¹³ To obtain objective data on this essential question, it seemed evident that a special study of the legibility of the L.C. cards and their reproductions would be necessary.

The Experimental Design

With the cooperation of the cataloging department of the Columbia University Libraries,¹⁴ a controlled experiment was planned and conducted to test the hypothesis that, in estimating costs of use by catalogers, an approximate equivalence in legibility of the printed L.C. cards and their bound reproductions, as far as this factor may affect working efficiency, legitimately can be assumed.

Twenty-four professional catalogers, all but three¹⁵ employed in the Columbia University Libraries, served as subjects for this

experiment. Three age groups, with eight subjects in each, were formed as follows: Group I, under thirty-five years; Group II, between thirty-five and forty-five years; Group III, over forty-five years. All but two of the subjects were women. Thirteen of the total group wore eyeglasses, three from Group I, and five each from Groups II and III. The amount of correction increased with the age group; there were no large corrections in Group I and no small ones in Group III.

Each group of subjects worked during four experimental sessions held on four successive mornings, each group beginning work on the same day of the week in successive weeks. Each work period started at 9 A.M. and continued until all subjects had completed the experimental task for that session. A large classroom in the School of Library Service, with individual work desks at which the subjects were seated according to a predetermined order, was used for the experiment. Blinds were drawn and indirect, overhead lighting was used to keep the illumination as constant as possible. The same experimenter served throughout all sessions. Each subject was provided with an individual desk clock and asked to record the time to the nearest minute in a designated space at the bottom of certain worksheets when a particular unit of work was completed.

The experimental task was so simple for these subjects that no practice period was needed. It consisted of transcribing a given amount of material from L.C. cards and from reproductions of these cards on mimeographed worksheets 8½ by 11 inches in size. A set of thirty worksheets, corresponding either to an equal number of cards or to an equal number of reproductions, comprised each assignment. During two of his four experimental sessions, the subject worked with L.C. cards in two trays placed on his desk; at the other two, he was provided with

¹³ In a recent study of the cost of using microfilm for reproduction of periodical articles, Shaw remarks that while business and industry do not ordinarily confuse first cost with total cost when they are actually different, librarians sometimes appear to do so. In his study, Shaw is not concerned with reading time or errors in perception in the use of microfilm. He starts with the assumption that "the time consumed in the actual reading of a given periodical article is substantially the same for the original, for a photoprint copy, and for a microfilm copy which does not require reference back and forth and which is ready to read." Obviously, this assumption is open to question. Shaw, Ralph R. "Should Scientists Use Microfilm?" *Library Quarterly* 14:229-33, July 1944.

¹⁴ Althea Terry, head of the cataloging department, was most helpful in working out experimental arrangements. Ilse Bry, a member of the department, served as research assistant under the direction of the writer, selecting and preparing the materials, conducting the experimental sessions, and assisting in analysis of data. Mrs. David Maxfield, former Columbia cataloger, scored the worksheets, and Rodman Bassein, also a member of the department, performed the statistical calculations. Invaluable help was given by Edwin G. Boring in designing the experiment.

¹⁵ Graduate students in the Columbia School of Library Service, all of whom had had cataloging experience.

two bound volumes containing the reproductions to be transcribed. Colored markers in the trays indicated the cards to be used, so that no time would be lost in locating them. Colored markers also were placed in the volumes to indicate the pages on which the designated reproductions were to be found; a small right-angled strip of easily detachable adhesive tape was pasted just outside the upper left corner of each reproduction to be used.

On every worksheet was mimeographed some of the material to be found on the card or on the reproduction to which it corresponded. Lines drawn below blank spaces indicated blocks of material that were to be filled in with pencil from the text of the card. This completion technique was used as a means for estimating the accuracy and speed with which the subjects could read the text on the L.C. cards as compared with the reproductions. Since constant reference was to be made to the text, this method of measurement eliminated as possible complicating variables both memory and comprehension of the material read (the latter was important in the case of unfamiliar foreign language material). Differences in handwriting speed among subjects could be disregarded, since subjects were not to be compared with one another. The crucial factor was the difference between efficiency in transcribing material from the cards and from the reproductions. It was assumed that each individual's speed of handwriting would remain relatively constant throughout the experiment.

To avoid practice effects, the material as a whole was divided into two halves, Part A and Part B, and arranged so that subjects who worked with the A half of the card material would transcribe from the B half of the reproductions. The A and B halves were carefully selected to obtain equivalence of content. Every subject worked with one set of cards and one

set of reproductions printed in English and with one set of each in selected foreign languages. The assignments were rotated systematically within each group of eight subjects, so that each subject began with a different set of material. This was a further precaution against practice effects. Each of the three age groups completed the same assignments, arranged in the same order. There were eight different sets of thirty worksheets in use at each session, representing 120 cards and 120 reproductions. Eight card trays and eight bound volumes were used at each session. Every subject completed in all 120 worksheets, half from cards and half from reproductions, in the four sessions during which he worked. Each set of thirty worksheets was completed by four subjects in each age group and by twelve subjects in all. Table I shows the order and character of each work assignment for each subject in all three groups.

In planning this experiment, another parameter in addition to the age and language variables had to be considered. This was type size. The L.C. cards are set in 12-, 10-, and 8-point type which is reduced in the reproductions to about 7-, 6-, and 5-point, respectively. In order to give equal representation to the three sizes of type used on the L.C. cards, which for purposes of this study are designated as large, medium, and small, ten worksheets respectively in each assignment of thirty were set up to require transcription of each of the three sizes of type. Material set in one type size only was transcribed on each worksheet, so that time scores could be kept separately for each size.

In working out an order of occurrence within each set of thirty worksheets for the three sizes of type and for the six possible sequences (small-medium, small-large, medium-small, medium-large, large-small, and large-medium), it was important to avoid any regularly recurring rhythms to which

the subjects' eyes might adjust in a way that might mask the effects of size differences. At the same time, it was necessary to secure a pattern that could be controlled throughout the experiment by systematic variation. To achieve these objectives, a table of random sampling numbers¹⁶ was used as a basis for constructing a satisfactory pattern.

place throughout his four sessions. Since the purpose of the experiment was to test the legibility of the material, not the relative efficiency of the subjects, competitive strain was avoided by assigning numbers to conceal the subjects' identities. In order to avoid possible intergroup rivalries, the subjects were not told until afterwards that age

TABLE I
Basic Experimental Design Used for Groups I, II, and III*

Subjects	<i>Experimental Sessions</i>			
	1	2	3	4
1	R EA	C EB	R FA	C FB
2	C EB	R EA	C FB	R FA
3	R FB	C FA	R EB	C EA
4	C EA	R EB	C FA	R FB
5	R FA	C FB	R EA	C EB
6	C FB	R FA	C EB	R EA
7	R EB	C EA	R FB	C FA
8	C FA	R FB	C EA	R EB

* R = Reproductions
C = Cards

E = English
F = Foreign

A = Part A (1st half)
B = Part B (2nd half)

The term "random" in this connection means (1) ignorance and unpredictability and (2) the following of prescribed laws of probability so that every item occurs equally often and every sequence equally often in samples that are not too small. In preparing the work assignments, the basic pattern was varied in four ways and the same set of variations was used with all three age groups. Thus, every subject, in each of his four assignments, worked through a different order of unpredictable recurrences of type sizes and sequences. Yet every subject in the total group was exposed, with systematic rotation in the pattern order, to the same four variations of the basic random pattern.

The subjects drew lots at the start of the experiment to determine their seating arrangement and each person kept the same

differences were to be studied. It was implied that the composition of the three groups was a matter of administrative convenience. In all other respects the subjects had been thoroughly acquainted by the writer with the purpose and design of the study. They were interested in the problem and they were promised a report on the findings. In a group conference held before the experiment began, opportunity had been given for questions and suggestions. The importance of approaching the experiment in a scientific spirit of conscientious and unbiased participation was stressed. The writer has every reason to believe that the cooperation of the entire group throughout the experiment was excellent.

Before each work session began the experimenter arranged all material in proper order on the subjects' desks. The members of the group began work at the same time on a signal from the experimenter. Pre-

¹⁶ Kendall, M. G., and Smith, B. Babington. *Tables of Random Sampling Numbers*. (Tracts for Computers, No. 24.) Cambridge, University Press, 1939, p. 33.

liminary experimentation had indicated that each assignment would require about an hour to complete, but this time was exceeded considerably by some subjects with some assignments. To avoid undue fatigue, a rest period of ten minutes after the first hour was provided. After each assignment was completed, the subject was given a short questionnaire on which he was asked to record his subjective reactions toward certain aspects of the experimental experience. At the conclusion of the last session for each group, a somewhat longer questionnaire was administered to obtain attitudes and opinions on the relative legibility of the cards and their reproductions.

Selection of the Material

Criteria for selecting the 120 cards and their matching reproductions were set up to meet the objectives of the experiment and to conform to the requirements of the experimental design. The content of the cards, as a whole, was representative of the kinds of material generally found on L.C. cards. Small amounts of special kinds of material, such as periodicals, documents, and legal material, were also included. A list of types of material that might be more difficult than the average to transcribe accurately was prepared and the variety and frequency of this material was regulated throughout the selection. An attempt was made to select cards whose reproductions were apparently of fairly average quality, with a small percentage respectively of better than average and poorer than average.

The selection of the material for this experiment was a highly exacting and time-consuming procedure. Limitations of space preclude further description of the many minor variables that were considered and controlled. Great care was taken to insure representative sampling and as complete equivalence as possible of the A and B parts of the material. This was necessary in

order to control all variables which might in any way affect the validity of the final comparison of the cards and their reproductions with respect to legibility.

Scoring Procedures

A scoring key for each assignment was prepared by typing on a copy of each of the worksheets an accurate transcript of the material to be filled in by the subject. Tables were made showing the number of printed characters to be transcribed on each sheet. The papers were first scored for errors, each deviation from a character on the key being counted as one error. Omissions or additions of characters likewise were scored as errors. Time scores were calculated, with totals for the three type sizes in each assignment kept separately. Net output of work per minute for each subject was then calculated by subtracting the number of errors from the number of characters to be transcribed and dividing the result by the number of minutes required by the subject to complete the assignment.

Results

This experiment was designed to test the hypothesis that an approximate equivalence in legibility of the printed L.C. cards and their bound reproductions, as far as this factor may affect working efficiency of catalogers, legitimately can be assumed. In statistical terminology, we are testing the hypothesis that the two samples under consideration (L.C. cards and their reproductions) were "drawn at random from identical normal populations, *i.e.*, from normal populations with the same mean and same standard deviation," or, stated more simply, "that they were drawn from the *same* population, since populations with identical distributions may be considered as constituting a single population."¹⁷ Such

¹⁷ Lindquist, Everet F. *Statistical Analysis in Educational Research*. Boston, Houghton Mifflin, 1940, p. 56. (Also see p. 15.)

a hypothesis—that the true difference between these two samples is zero—is known as a “null hypothesis.” If, by statistical analysis of the experimental data, we find that the difference in mean achievement of catalogers working with these two samples is larger than could reasonably be attributed to fluctuations in random sampling, then we may reject the null hypothesis. Before we can say that this difference in mean achievement can be attributed to a difference in legibility, however, we must adequately have controlled other variables in the situation which might conceivably have caused the difference. The experimental design of this study was planned to control such variables. By statistical analysis of the data, the degree to which this control was exercised can be determined.

Equivalence of Parts A and B

The first factor to be considered is the equivalence of Parts A and B of the material. We must test the hypothesis that these parts were equally matched, *i.e.*, that there is no significant difference between the samples of cards selected respectively

to compose these two halves of the material. Tables II and III give the differences between the mean achievement of the catalogers on Parts A and B of the material in terms of errors and net output per minute. These differences are broken down into English and foreign language material and into the three age groups for both cards and reproductions. To determine the significance of the differences, the *t*-test was used.¹⁸ Various levels of significance for the obtained values of *t* were ascertained by use of Fisher and Yates's table.¹⁹ The level of significance indicates the number of chances in one hundred that the two samples were not drawn from the same or identical populations. When the value of *t* is greater than that required for a given level of significance, the null hypothesis may be considered disproved with a corresponding degree of confidence.

Following R. A. Fisher, statisticians in re-

¹⁸ Fisher, Ronald A. *The Design of Experiments*. 3d ed. London, Oliver & Boyd, 1942, p. 33-37. The formula used for calculating *t* can be found on p. 57 (formula 11) in Lindquist (*op. cit.*).

¹⁹ Fisher, Ronald A., and Yates, Frank. *Statistical Tables for Biological, Medical, and Agricultural Research*. 2d ed. London, Oliver & Boyd, 1943. Table III, p. 30.

TABLE II
Parts A and B—Errors

Measures	English					
	Cards			Reproductions		
	I	Group II	III	I	Group II	III
Mean A	9	4	19	196	9	437
Mean B	13	3	22	71	63	208
Difference	-4	1	-3	125	-54	229
Significance*						
Measures	Foreign					
	Cards			Reproductions		
	I	Group II	III	I	Group II	III
Mean A	47	40	11	308	47	749
Mean B	68	11	41	32	92	394
Difference	-21	29	-30	276	-45	355
Significance*						

* None of the differences is statistically significant.

TABLE III
Parts A and B—Net Output per Minute

Measures	<i>English</i>					
	Cards			Reproductions		
	I	Group II	III	I	Group II	III
Mean A	96	86	85	67	72	64
Mean B	76	81	79	83	73	68
Difference	20	5	6	-16	-1	-4
Significance*	<i>b</i>					

Measures	<i>Foreign</i>					
	Cards			Reproductions		
	I	Group II	III	I	Group II	III
Mean A	81	67	61	52	55	43
Mean B	61	66	61	72	57	54
Difference	20	1	0	-20	-2	-11
Significance*						

* *b* = from 95 to 98 chances in 100 that the means are significantly different.

cent years have tended to accept the 5-per cent and 1-per cent levels of significance. The 5-per cent level means that there is one chance in twenty that the two samples whose difference is being considered have been drawn from the same population, *i.e.*, there is one chance in twenty that the difference is not significant and therefore nineteen chances in twenty, or ninety-five in one hundred, that the difference is significant. Fisher's terms, however, are confusing because the difference becomes more significant as the stated probability (that there is *no* significant difference) becomes smaller. In this paper, the 95-per cent level of the significance of a difference has been used to designate what in Fisher's terminology is called the 5-per cent level; similarly, the 99-per cent level is what Fisher calls the 1-per cent level, and so on.

In Tables II and III (and the following tables), the significance of each-difference is indicated in the last row, for various levels of significance, by the following set of symbols:

no symbol = less than 90 chances in 100 that the means are significantly different

a = from 90 to 95 chances in 100
b = from 95 to 98 chances in 100
c = from 98 to 99 chances in 100
d = from 99 to 99.9 chances in 100 (or 999 in 1000)
e = greater than 999 chances in 1000

To save space, the calculated values for *t* are not given in the tables. Means were worked out to two decimals but, again to save space, the decimals were dropped after reduction.

So far as errors are concerned, Table II shows that no consistent pattern of differences between Parts A and B, as dependent on their content, can be discerned. While some relatively large differences²⁰ appear, these depend upon the age group, the language, and whether the work sample is a reproduction. The equivalence of the A and B parts of the material was sub-

²⁰ Considerable variability, both in number of errors and in net output per minute, was found among the individual subjects. This was not a function of the age group. The smallest number of errors made by any individual in the four work assignments was 32, the largest 3901. In working with the reproductions, some subjects omitted whole sections of the small type material in some samples because they could not read them. These omissions were counted as errors. The best score in terms of net output per minute was 100.4 characters, the poorest was 46.6. The rank correlation between errors and net output per minute was .39 for the twenty-four subjects.

stantiated by the small *t* values, which, as the absence of significance symbols in the last row of the table shows, were in no case sufficiently greater than zero to indicate as many as 90 chances in 100 that the A and B parts were drawn from different populations.

A similar conclusion can be drawn on the basis of the data for net output per minute. Table III indicates that the B cards are in general somewhat less rapidly handled than the A cards, but this difference

the foreign language material. Here we are interested in the question, if differences are found, as to whether these differences show a consistent pattern in both the cards and the reproductions. If they do, we may assume that the language differences are not the cause of any differences in efficiency that may be found when the catalogers' work with the cards is compared with that done with the reproductions. It is apparent from these tables that the English material is consistently somewhat easier to transcribe

TABLE IV
Languages E and F—Errors

Measures	Group I			Cards Group II			Group III		
	Type Size			Type Size			Type Size		
	L	M	S	L	M	S	L	M	S
Mean E	2	2	7	1	2	1	3	4	13
Mean F	6	14	37	2	3	21	2	5	19
Difference	-4	-12	-30	-1	-1	-20	1	-1	-6
Significance*				<i>a</i>		<i>b</i>			

Measures	Group I			Reproductions Group II			Group III		
	Type Size			Type Size			Type Size		
	L	M	S	L	M	S	L	M	S
Mean E	3	34	97	1	29	5	2	16	304
Mean F	12	36	123	15	14	40	20	174	378
Difference	-9	-2	-26	-14	15	-35	-18	-158	-74
Significance*						<i>b</i>	<i>b</i>		

* *a* = from 90 to 95 chances in 100, *b* from 95 to 98 chances in 100 that the means are significantly different.

cannot be due to the content of the cards since the situation is reversed in the reproductions. The *t* values, although in some cases higher than they were for the errors, show only one difference above the 90-per cent level of significance. These findings support the hypothesis that there is no significant difference between the A and B parts of the material and make it possible to combine them in further analyses of the data.

Tables IV and V are concerned with the differences in difficulty of the English versus

than the foreign material, whether the subject is working from the cards or from the reproductions.

In every case, the catalogers worked more slowly with the foreign material, as is shown in Table V. This held true in all three sizes of type. Seven of the nine differences in the card material lie above the 90-per cent level of significance, as do six of the nine differences in the reproductions. Seven of these thirteen differences lie above the 99-per cent level. Although the catalogers worked more slowly with the foreign

samples, they nevertheless tended to make more errors with them, as appears in all but two of the comparisons shown in Table IV. Only four of these differences, however, lie above the 90-per cent level of significance.

vary directly with age since the youngest and the oldest catalogers tended to make more errors than those in Group II. Again, the great variability among individuals in all three age groups accounts for the fact that large differences found in some in-

TABLE V
Languages E and F—Net Output per Minute

Measures	Group I			<i>Cards</i> Group II			Group III		
	Type Size			Type Size			Type Size		
	L	M	S	L	M	S	L	M	S
Mean E	71	95	83	76	90	80	75	88	79
Mean F	63	75	70	56	71	66	57	64	60
Difference	8	20	13	20	19	14	18	24	19
Significance*		<i>b</i>		<i>d</i>	<i>d</i>	<i>d</i>	<i>d</i>	<i>e</i>	<i>d</i>

Measures	Group I			<i>Reproductions</i> Group II			Group III		
	Type Size			Type Size			Type Size		
	L	M	S	L	M	S	L	M	S
Mean E	61	85	72	57	80	69	59	73	61
Mean F	51	67	61	54	60	54	44	54	47
Difference	10	18	11	3	20	15	15	19	14
Significance*		<i>a</i>			<i>d</i>	<i>c</i>	<i>b</i>	<i>c</i>	<i>a</i>

* *a* = from 90 to 95 chances in 100, *b* from 95 to 98 chances in 100, *c* from 98 to 99 chances in 100, *d* = 999 chances in 1000, and *e* = greater than 999 chances in 1000 that the means are significantly different.

Effect of Age

The question as to whether there are any significant differences in working efficiency on the experimental task among the three age groups may be answered by reference to Tables VI and VII. In general, the older catalogers (Group III) worked a little more slowly than those in Group II, while the middle group in turn tended to work a little less rapidly than the youngest (Group I). While these differences, as shown in Table VII, are fairly consistent, they are relatively small and only two of the thirty-six differences computed reach the 90-per cent level of significance. Table VI shows slight differences on the average in the number of errors made by the different age groups, but accuracy does not

stances between the group means are for the most part not significant statistically. Only one difference was found between Groups I and III which exceeded the 90-per cent level in the twenty-four comparisons made. Between Groups I and II, four of the twenty-four differences were significant with three favoring the middle group; three significant differences were found between Groups II and III which also favored the middle group.

If, in spite of the great variability within our small groups, there appears to be some tendency for accuracy to increase with age, then it would seem reasonable to suppose that two factors are acting simultaneously in opposite directions. It is possible that the younger group has the advantage of

TABLE VI
Age Groups I, II, and III—Errors

Measures	Cards						Reproductions					
	English			Foreign			English			Foreign		
	Type Size			Type Size			Type Size			Type Size		
	L	M	S	L	M	S	L	M	S	L	M	S
Mean I	2	2	7	6	14	37	3	34	97	12	35	123
Mean II	1	2	1	1	3	21	1	29	5	15	14	40
Mean III	3	4	13	2	5	19	2	16	304	20	174	378
Diff. I—II Significance*	1 <i>b</i>	0	6 <i>e</i>	5 <i>a</i>	11	16	2	5	92	-3	21	83
Diff. II—III Significance*	-2 <i>a</i>	-2	-12 <i>b</i>	-1	-2	2	-1	13	-299	-5	-160	-338 <i>a</i>
Diff. I—III Significance*	-1	-2	-6	4	9	18	1	18	-207	-8	-139	-255

* *a*=from 90 to 95 chances in 100, *b* from 95 to 98 chances in 100, and *e*=greater than 999 chances in 1000 that the means are significantly different.

better visual acuity, an advantage which diminishes with age but is offset among the older catalogers by their experience in dealing with this sort of material. In any case, the achievement of the three age groups is sufficiently uniform from a statistical standpoint to permit us to combine them in our final analysis of the differences between cards and reproductions.

Effect of Type Size

We are now ready to examine the differences among the three sizes of type. These differences are exhibited in Tables VIII and IX. With respect to errors, there are no significant differences between the large and medium type. Three of the twelve differences between the medium and the small type are significant above the

TABLE VII
Age Groups I, II, and III—Net Output per Minute

Measures	Cards						Reproductions					
	English			Foreign			English			Foreign		
	Type Size			Type Size			Type Size			Type Size		
	L	M	S	L	M	S	L	M	S	L	M	S
Mean I	71	95	83	63	74	70	61	85	72	51	67	61
Mean II	76	90	80	56	70	66	57	80	69	54	60	54
Mean III	75	88	79	57	64	60	59	73	61	44	54	47
Diff. I—II Significance*	-5	5	3	7 <i>a</i>	4	4	4	5	3	-3	7	7
Diff. II—III Significance*	1	2	1	-1	6	6	-2	7	8	10	6	7
Diff. I—III Significance*	-4	7	4	6	10	10	2	12	11	7	13	14 <i>a</i>

* *a*=from 90 to 95 chances in 100 that the means are significantly different.

TABLE VIII
Type Sizes L, M, and S—Errors

Measures	Cards						Reproductions					
	English			Foreign			English			Foreign		
	I	Group II	III	I	Group II	III	I	Group II	III	I	Group II	III
Mean L	2	1	2	6	2	2	3	1	2	12	15	20
Mean M	2	2	4	14	3	5	34	29	16	36	14	174
Mean S	7	2	13	37	21	19	97	5	304	123	40	378
Diff. L—M Significance*	0	-1	-2	-8	-1	-3	-31	-28	-14	-24	1	-154
Diff. M—S Significance*	-5 <i>e</i>	0	-9	-23	-18 <i>b</i>	-14 <i>a</i>	-63	24	-288	-87	-26	-204
Diff. L—S Significance*	-5 <i>e</i>	-1	-11 <i>a</i>	-31	-19 <i>b</i>	-17 <i>b</i>	-94	-4 <i>a</i>	-302	-111	-25	-358 <i>a</i>

* *a* = from 90 to 95 chances in 100, *b* from 95 to 98 chances in 100, and *e* = greater than 999 chances in 1000 that the means are significantly different.

90-per cent level. When the large and the small type are compared, half of the differences are statistically significant. In all cases where differences are significant, more errors occur when the catalogers are working with the small type. Although fewer errors are made with the large type, the catalogers seem to have worked more slowly when transcribing the large-type material. When the net output per minute for the large and the medium-sized type is compared, the results show a difference in

favor of the medium size that exceeds the 90-per cent level of significance in eight of the twelve computations. Likewise, the differences between the large and the small type favor the latter, although only two of these differences are statistically significant.

The explanation for this finding seems to lie in the probability that because greater difficulty was experienced in reading the smaller-type samples that preceded the large-type blocks, the subjects unconsciously

TABLE IX
Type Sizes L, M, and S—Net Output per Minute

Measures	Cards						Reproductions					
	English			Foreign			English			Foreign		
	I	Group II	III	I	Group II	III	I	Group II	III	I	Group II	III
Mean L	71	76	75	63	56	57	61	57	59	51	54	44
Mean M	95	90	88	75	71	64	85	80	73	67	60	54
Mean S	83	80	79	70	66	60	72	69	61	61	54	47
Diff. L—M Significance*	-24 <i>c</i>	-14 <i>b</i>	-13 <i>b</i>	-12	-15 <i>c</i>	-7	-24 <i>d</i>	-23 <i>d</i>	-14	-16 <i>a</i>	-6	-10 <i>a</i>
Diff. M—S Significance*	12	10	9	5	5	4	13	11 <i>a</i>	12	6	6	7
Diff. L—S Significance*	-11	-4	-4	-7	-10 <i>c</i>	-3	-11	-12 <i>a</i>	-2	-10	0	3

* *a* = from 90 to 95 chances in 100, *b* = from 95 to 98 chances in 100, and *d* = 999 chances in 1000 that the means are significantly different.

relaxed and slowed down when they came to the selections that were easier to read. At the same time, because the large-type samples were more legible, the catalogers were able to transcribe them with greater accuracy. In the final comparisons between

loggers made many more errors, in both the English and foreign material, when working with the reproductions than when transcribing from the cards. The only exception was in the case of the large-sized type where no difference was found. All of the differences

TABLE X
Cards and Reproductions (Languages E and F)—Errors

Measures	English Type Size			Foreign Type Size		
	L	M	S	L	M	S
Mean C	2	3	8	3	8	26
Mean R	2	26	135	16	75	180
Difference	0	-23	-127	-13	-67	-154
Significance*		c	a	c	a	b

* a = from 90 to 95 chances in 100, b from 95 to 98 chances in 100, c from 98 to 99 chances in 100 that the means are significantly different.

TABLE XI
Cards and Reproductions (Languages E and F)—
Net Output per Minute

Measures	English Type Size			Foreign Type Size		
	L	M	S	L	M	S
Mean C	74	91	81	59	70	65
Mean R	59	80	67	50	61	54
Difference	15	11	14	9	9	11
Significance*	e	d	d	c	b	c

* b = from 95 to 98 chances in 100, c from 98 to 99 chances in 100, d = 999 chances in 1000 that the means are significantly different.

cards and reproductions, the three type sizes were analyzed separately, since it seemed clear that they were related to the legibility factor.

Cards Versus Reproductions

Tables X through XIII present these final analyses on the differences between the cards and the reproductions. A preliminary breakdown was made, with the English and foreign languages treated separately. Table X shows that the cata-

loggers made many more errors, in both the English and foreign material, when working with the reproductions than when transcribing from the cards. The only exception was in the case of the large-sized type where no difference was found. All of the differences

are above the 90-per cent level of significance. That the catalogers also worked much more slowly with the reproductions than with the cards in both English and foreign material and in all three type sizes is demonstrated in Table XI. All of these differences exceed the 95-per cent level of significance. While for both cards and reproductions, in both English and foreign material, the errors show a consistent increase from the large to the small-sized

type, the net output per minute is greatest for the medium-sized type and smallest for the large type. These findings are consistent with those shown in Tables VIII and IX.

Tables XII and XIII present the final data on the efficiency of the whole group of subjects in transcribing from the cards

consistently as the type sizes became smaller.

Use of the reproductions cut down considerably the efficiency of the group in terms of the number of characters transcribed. In all three type sizes the net output per minute was less for the reproductions than for the cards. The differences were all

TABLE XII
Cards and Reproductions (Type Sizes)—Errors

Measures	Type Size			Total
	L	M	S	
Mean C	2	5	17	24
Mean R	9	51	158	218
Difference	-7	-46	-141	-194
Significance*	<i>c</i>	<i>c</i>	<i>d</i>	<i>d</i>

* *c* = from 98 to 99 chances in 100, *d* = 999 chances in 1000 that the means are significantly different.

TABLE XIII
Cards and Reproductions (Type Sizes)—
Net Output per Minute

Measures	Type Size			Total
	L	M	S	
Mean C	66	80	73	73
Mean R	54	70	61	62
Difference	12	10	12	11
Significance*	<i>e</i>	<i>d</i>	<i>e</i>	<i>e</i>

* *d* = 999 chances in 1000, *e* = greater than 999 chances in 1000 that the means are significantly different.

and the reproductions when the language differences are ignored. As shown in Table XII, errors occurred with much less frequency when the catalogers were transcribing from the cards than when they were working with the reproductions of the same material in the bound volumes. The differences in all three sizes of type were above the 98-per cent level of significance. In the small-sized type an even higher level of significance was exceeded, as was the case for the total computation when all three type sizes were combined. In both cards and reproductions, the errors increased

above the 99-per cent level of significance. These results are shown in Table XIII. Here, again, it is evident that the catalogers as a group worked faster with the medium-sized type than with the large or the small and faster with the small than with the large type. As explained above, this was probably due to the fact that the subjects slowed down on the large-type material while recovering from strain induced by difficulty experienced in working with the smaller print.

From the standpoint of use cost of the cards as compared with the reproductions,

the findings of this experiment indicate that the net output per minute of these catalogers when working with the reproductions is about 85 per cent of the net output for the cards. This means that the difference in output is sufficiently great to require seven catalogers working with the reproductions to do the work of six catalogers using the cards. These seven catalogers will, with the reproductions, make about nine times as many errors as they would with the cards. The errors will occur in all sizes of type: approximately 73 per cent in the small type, 23 per cent in the medium-sized type, and the remaining 4 per cent in the large type. Errors will be more numerous and output less when the catalogers are working with foreign language as compared with English material.

Subjects' Introspections

Space permits only a brief summarization of the reactions of the catalogers who participated in this experiment. Before beginning work on the first assignment the subjects were asked to check on a five-point scale their attitude toward a decision to substitute the bound volumes for the depository catalog in the cataloging department of their library. After they had completed two work sessions they were asked to check another copy of this scale; at the end of the last session, they again indicated their attitude. The following tabulation of the number of subjects checking each point on the scale shows how these attitudes shifted:

<i>Attitude</i>	<i>Before 1st Session</i>	<i>After 2d</i>	<i>After 4th</i>
Strongly approve	1	1	0
Approve	4	6	7
Undecided	12	3	1
Disapprove	4	7	4
Strongly disapprove	3	7	12

At the end of the experiment sixteen of the twenty-four catalogers were convinced

that it would be a mistake to substitute the reproductions for the cards. The consensus of this majority was that use of the reproductions would undermine the validity of their professional work since they could never be sure, unless they had a given book in hand, whether or not they were reading the text on the card correctly. Two of this group said that they would prefer to give up cataloging entirely rather than experience the feelings of uncertainty, frustration, and strain which use of the reproductions engendered. Many of the participants reported irritability, annoyance, and exasperation at the inadequacy of the bound volumes as a bibliographical tool.

With respect to the question as to whether a reading glass would obviate some of the difficulties encountered in working with the reproductions, opinion was divided. Some thought such a verification device would clear up uncertainties, but others felt that there were many defects in the type of the reproductions that could not be overcome by magnification of the text and that the reduction in type size was only one, and perhaps not the most important, factor of illegibility. Many commented on the fact that the experimental sessions were much longer than periods of time spent in working with the reproductions for professional purposes. They also pointed out that the experimental task differed considerably from the actual work of a cataloger.²¹

Those who approved the proposed substitution of the bound volumes for the depository catalog felt that the disadvantages of the reduced type were outweighed by the greater convenience of use, the ease with which material may be located, saving of space, and "up-to-dateness" of the entries.

²¹ The work assignments had to be sufficiently long to insure reliability of the data. The experimental task was designed to measure the *legibility* of the material, not the use to which it would be converted by professional catalogers.

It is evident that the difference in opinion with regard to the usefulness of the reproductions of the Library of Congress cards found among the Columbia catalogers after they had participated in this experiment reflects the divergence of views in the field. This finding gives further indication of the need for objective data as an indispensable aid in understanding the factors involved in this problem.

Conclusion

The results of this study show that in estimating costs of use by catalogers of the printed L.C. cards, as compared with the bound reproductions, the legibility factor must be taken into account. Use of the printed cards affords greater working efficiency in that catalogers work more rapidly and more accurately with the cards than with the reproductions.

In a given amount of time the catalogers transcribed from the reproductions only about 85 per cent as much material as from the cards. In other words, in respect of output alone, it would take seven catalogers working with the reproductions to do the work of six with the cards. That, however, is not all. The seven catalogers working with the reproductions will make about nine times as many errors as the six catalogers working with the cards. Most of the errors (73 per cent) will occur in the small type, some (23 per cent) in the medium-sized type, and a few (4 per cent) in the large type.

In all cases transcription was slower and errors were more numerous in the foreign languages as compared with English.

These differences held for subjects in all three age groups.

In general, the subjects began the experiment without great prejudice for or against the reproductions. Five of them expected to find the reproductions more satisfactory than the cards, seven of them

expected the reproductions to be less satisfactory than the cards, the remaining twelve were undecided. At the end of the experiment, every subject but one had made up his mind. Seven were impressed by the greater convenience of the bound volumes, one was undecided, whereas the remaining sixteen disapproved or strongly disapproved the reproductions because the illegibility of the material undermined confidence in the accuracy of their work. This led to feelings of frustration, irritation, and strain.

It will be necessary to study carefully the actual extent of use of the depository catalog before it will be possible to say whether the convenience, the lower first costs, and the lower cost of maintenance of the bound volumes justifies the inefficiencies which use of the reproductions entails.²²

Since inequalities in paper and printing of the original printed L.C. cards exist, apparently the only way to obtain more legible reproductions is to make them larger, *i.e.*, to use a smaller reduction in type size. It is, of course, possible to reset the material in type but this would be much more expensive and would probably introduce errors in the text. Elimination of white space in the reproductions would allow some enlargement without increasing the bulk of the volumes and this advantage might easily be worth the extra labor involved in preparing the material for reproduction.

It is to be hoped that in planning for supplements to the bound volumes of the Library of Congress catalog or for reproductions of similar catalogs (*e.g.*, the Catalog of the British Museum) this question of the size of reproduction can be considered more carefully in view of the disadvantage that too great reduction in type size is now known to entail.

²² A study of this variable as it affects the situation in the Columbia University Libraries is now in progress. This is "Use of the Depository Catalog in a University Library" by Frances Munson, a master's thesis in the School of Library Service, Columbia University.