Information Processing & Management Vol. 20, No. 1-2, pp. 229-237, 1984
Printed in Great Britain.

0306-4573/84 $3.00 + .00
Pergamon Press Ltd.

EXPERIMENTAL AND QUASI-EXPERIMENTAL DESIGNS FOR RESEARCH IN INFORMATION SCIENCE

DAVID F. HAAS and DONALD H. KRAFT
Department of Computer Science, Louisiana State University, Baton Rouge, LA 70803, U.S.A.

Abstract—This is a paper about research designs in information science. We look at a sample of current research and compare its designs with an abstract ideal of experimental research design to see how closely they approximate it. We then consider ways in which research in our field might be brought closer to the ideal. We do this because we believe that experimental and quasi-experimental designs offer unique advantages over other research designs, especially in the production of knowledge that can be applied to the solution of practical problems in information and in software science.

1. WHY DO EXPERIMENTS?

Information science has been working to define its scientific basis. This is may be seen in the efforts of the Special Interest Group on Foundations of Information Science (SIG/FIS) and in such publications as FELDMAN[4]. In this paper we hope to contribute to this attempt by examining the methodological bases of empirical research in information science. What do we mean by "empirical?" In an empirical study, the researcher gathers and analyzes data from the "real" world to test the validity of some hypothesis about it. This means that strictly theoretical investigations are not empirical. Studying the behavior of a mathematical model of a process without testing the model against data also falls outside of the realm of the empirical. For, the model is simply a mathematical expression of a theory.

Experimental research is one type of empirical research. Experimental designs are to be preferred to other research designs when the goal of the research is to produce reliable *causal* knowledge. That is, when we want to be reasonably confident that some event is caused by another event and not merely correlated with it, the best way to test our belief is by experimentation. The notion of "cause" is certainly one of the murkiest in science. HUME[11] showed long ago that the term has no clear, empirical referent beyond a constant conjunction of two events. We can never show conclusively that when one event follows another, the latter is produced by the former and hence will always follow it unless something else intervenes.

Nevertheless, if science is to produce knowledge that can be used technically to solve practical problems, that knowledge must take the form of causal assertions. On the practical level, we must act as if we could say that if we produced some event, it would in turn produce some other event or state. Thus, we say that innoculation of our children against polio causes them to acquire an immunity to the disease or that the introduction of an online bibliographic searching facility causes literature searches to be performed more easily or quickly.

When we want knowledge that can be used in this way, experimentation is the best way to test our beliefs. For, experimentation differs from other types of research precisely in the fact that in an experiment the researcher manipulates the "independent variable" (or treatment) under controlled conditions and then observes the behavior of the "dependent variable" (or outcome). Thus, he can test the hypothesis that a change in the treatment will bring about a change in the outcome. The recent emphasis on "experimental" computer science shows the importance of the experimental form of research[4]. We hope to offer some definitions and standards suitable for both information and software science.

Let us be more precise about what we mean by "controlled conditions." To be an experiment, a study must have at least the following characteristics[1]:

(1) *A treatment.* The treatment is the independent variable. It must be produced or manipulated by the experimenter.

(2) *An outcome measure.* This is the measurement of the dependent variable after the treatment has occurred.

(3) *A comparison measure.* This is a measure from which change in the dependent variable can be inferred. Comparison measures are of two types:

(a) *Comparisons across groups.* If there is more than one experimental group and if a different level of the treatment is given to each of them, then they may be compared to see whether the different treatments produced different outcomes.

(b) *Comparisons across time.* If the dependent variable is measured both before and after the treatment, then change may be inferred. Whether or not the change may validly be attributed to the treatment depends on whether plausible, alternative explanations have been ruled out.

(4) *Units of assignment.* Some units, e.g. individual people or groups, must be assigned to treatments in a controlled way. For example, in an an experimental study of the behavior of individuals in bibliographic searching under two different sets of conditions, the conditions would be the treatments, and the individuals would be the units of assignment. That is, they would be assigned to the treatments. If the units are *randomly assigned* to treatments, we speak of a "randomized experiment". Randomization, when it is feasible, is the preferred method for eliminating most alternatives to the experimental treatment as explanations of observed changes in the outcome. When randomization is not possible, we will call the study a "quasi-experiment".

2. THREATS TO THE VALIDITY OF CONCLUSIONS

Most writers distinguish between two types of validity, internal and external ([12], p. 301). Findings from an experiment are internally valid if we may properly conclude that the treatment did indeed cause the outcome *within the experimental setting.* The conclusions are externally valid if they may properly be generalized to other times or places. A finding might be internally but not externally valid if, for instance, the experimental arrangements themselves caused the subjects to attend to certain aspects of the treatment rather than to others. Thus, subjects have been known to figure out what the experimenter wants and to try to give it to him. Threats to internal validity include the following:

(1) *History,* the specific events that occurred during the study, other than the change in the independent variable, that might account for changes in the dependent variable.

(2) *Maturation,* or processes within the experimental subjects or respondents that operate as a function of the passage of time. For instance, subjects might get hungry or tired. In a study that runs for weeks or months, maturation can be a very serious problem. This would not be uncommon for instance, in studies of the use of query languages where the subjects first have to learn the languages.

(3) *Testing,* which is the effect of taking one test on the scores obtained on a second test. For, instance, in an ordinary classroom, students often try to use the first test given by an instructor to learn what kinds of questions he asks. So, their grades on later tests are due in part to the degree of their success in this effort. Thus, the later tests are not pure measures of their success in learning the course material.

(4) *Instrumentation,* in which changes in the calibration of a measuring instrument or changes in observers or scorers may produce changes in obtained measurements. For instance, when human scorers are used to rate computer programs for "goodness of programming style" the scorers' criteria may change in subtle ways as they go through the process of scoring. Even when an attempt is made to specify clear criteria, problems arise in their application ([6], pp. 21–22).

(5) *Statistical regression,* operating where groups have been selected on the basis of extreme scores. If a group is selected from a population on the basis of extreme scores on some measure, then if the same measure is repeated a second time, the scores of the group will on the average move closer to the mean of the population from which the group was

chosen. This is a widespread problem in psychology where one often wants to create groups whose members have extreme scores on self esteem, intelligence or some other characteristic.

(6) *Selection* biases, that is selection practices that produce groups whose scores on the dependent variable may be different, not because of the effect of the experimental treatment, but because of the composition of the groups.

(7) *Experimental mortality* or differential loss of respondents from comparison groups.

(8) *Selection-maturation interaction.* This occurs when subjects in different experimental groups change or mature at different rates because of selection biases. If this occurs, a difference between the two groups not due to the experimental treatment may appear even though there is no "main effect" of maturation. (The term "main effect" here has the technical sense given to it in analysis of variance (ANOVA). For a discussion of the distinction between a main effect and an interaction effect see [8], pp. 387–390.

In additon to these threats to internal validity, there are several threats to external validity, which are:

(1) *The reactivity of a test*, that is, the degree to which it causes those taking the test to behave differently from those who do not take it.

(2) *Interaction of selection and the treatment.* Just as maturation can interact with selection if the latter is not random, so can the treatment.

(3) *Reactive effects of the experimental arrangements.* Just as measure of the dependent variable may be reactive, so may other aspects of the experimental situation. Thus, for instance subjects may behave in special ways simply because they know that they are experimental subjects. This was, for instance, the basis of the famous "Hawthorne effect"[9].

(4) *Multiple treatment interference.* This occurs when the experimental treatment has a sensitizing effect on the experimental subjects. For instance, in an experiment requiring the subjects to perform a series of tasks, the order in which the tasks are presented may effect the subjects' relative performance on each of them if the skills learned in one task can be used in the next.

3. ELIMINATING THREATS TO VALIDITY

The goal of experimental design is to reduce or eliminate threats to the validity of experimental findings. Since randomization plays such a large role in this effort, and since it is so rare in research in information science, let us consider it first. We will consider it in the context of a very simple experimental design, one in which there is only a single treatment and in which there are only two groups, and "experimental group" (which receives the treatment) and a "control group" (which does not receive the treatment). Subjects are randomly assigned to the groups. Such a design may be diagrammed as follows, using Cook and Cambell's[2] conventions:

Experimental group:	R	X	O
Control group:	R		O

A Posttest Only Control Group Design

In the above diagram, R indicates random assignment, X indicates an experimental treatment, and O indicates an observation or measurement of the outcome or dependent variable. The design is a "posttest only" design because the outcome variable is measured only after the treatment. A measure taken before the treatment would be a "pretest". The use of the posttest only design is possible because randomization allows us to assume that there was no systematic difference between the groups before the treatment. Let us look more closely to see how randomization deals with most threats to the validity of conclusions drawn from an experiment of this design.

History is effectively dealt with here because any events in the world that occurred during the study must have affected both groups. Note also that any interaction of history with the selection of the subjects is taken care of by random assignment. The probability that results are affected by such interactions may be made arbitrarily small by increasing the sizes of the groups. *Maturation* is eliminated in the same way. *Instrumentation* may or may not have been eliminated here. If the outcome measure comes from a rating by observers and if all of the members of one group were tested before the members of the other group, then instrumentation might well still confound the results. If, on the other hand, the groups were run at the same time, or if individuals were tested in random order, then this threat is also eliminated. *Statistical regression* is a process that operates only when groups are selected on the basis of extreme scores. As mentioned above, this practice is quite common in psychology, but we have not observed it in our survey of research in information science. So, we will not discuss it further except to point out that it is entirely disposed of by random assignment. *Selection* biases that might affect the results by influencing the composition of the experimental groups are eliminated here, too. Of course, random assignment does not eliminate selection biases absolutely, but it eliminates them with a known probability which may in principle be made arbitrarily small.

Random assignment does not eliminate *experimental mortality* (*the loss of subjects from experimental* groups) or any of the threats to the *external* validity of the findings. This is because the random assignment of subjects to experimental conditions does not assure that the subjects as a group are representative of some larger population. Representativeness may be achieved by random *sampling* (the random selection of subjects from a population). This should not be confused with random assignment. We should note finally that random assignment, by equating the experimental and control groups initially, allows us to attribute differences between them after treatment to the experimental variable even though we administered no pretest. This is important because the pretest can, by sensitizing the subjects to its contents, serve as a confounding factor.

4. SOME REPRESENTATIVE RESEARCH IN INFORMATION SCIENCE

With the preceding in mind, we now turn to an examination of a representative sample of published works in information science to see how closely the model of experimentation presented here approximates the kind of work actually done in the discipline. In the examination, it will be assumed that scholars reporting the results of research will emphasize in their reports the aspects of the research designs that render their findings convincing. Moreover, we will assume that scholars will generally use rhetoric appropriate to the cases they are making and will not waste scarce journal space presenting irrelevant arguments. On the other hand, the failure to present crucial arguments would render the cases unconvincing, and therefore we expect that papers embodying such failures would be rejected, at least by the more prestigious journals. Finally, a prestigious journal's editors and reviewers may be assumed to judge articles according to the most widely accepted standards of the discipline. So, papers will be accepted only if they successfully present arguments regarded as appropriate by the standards of the discipline. Therefore, one may discover the methodological standards of a discipline by examining the persuasive rhetoric used in its most prestigious journals.

Table 1. Characteristics of articles surveyed

Independent Variable is Manipulated		Variables are Analytically General		Alternative Explanations Are Controlled	
yes	no	yes	no	yes	no
4	23	18	9	8	19

Note. The numbers in the table do not total to 31, the number of articles surveyed. The reason for this is that some of the papers were strictly tests of methods. There really were no "independent" or "dependent" variables because there were no hypotheses. For such studies, the categories shown here were not really meaningful.

A sample of the persuasive rhetoric used in information science was obtained from recent issues of three journals:

Journal of the American Society for Information Science,
Information Processing and Management,
Journal of Library Research.

The issues examined were all from 1981 or 1982, and articles reporting the results of empirical research were chosen randomly without replacement from each title. Strictly theoretical or methodological papers were rejected as were review articles, opinion pieces and papers reporting the results of modelling studies. In all we examined thirty-one articles. For a list of the papers surveyed, see the Appendix.

The results of our research are shown in Table 1 above. The table shows quite clearly that research in information science is not generally carried in accord with the standards of the model presented above. Only four of the studies employed any manipulation of the independent variable by the researcher. Moreover, in three quarters of the studies, no attempt was made to eliminate alternative explanations of the findings either by randomization or by statistical controls. This finding is quite surprising and, on the surface, rather dismaying. It means that the vast majority of studies in information science provide no useful test of any theoretical hypothesis. Let us look more closely at some of the studies to see just what kinds of research information scientists really do.

Methodological research

One kind of research that is quite common in information science is the test of a technique or method for doing some practical task. For instance, ZAMORA *et al.*[31] report a test of a method called "trigram analysis" for checking a text for spelling errors. (*Note.* Numbers in parentheses refer to the list in the Appendix of the articles surveyed for this paper.) Zamora's study is actually an exceptionally good one because he and his colleagues did manipulate one of their independent variables (subject matter of text). However, they made no attempt to eliminate alternative explanations for their findings. The question we must ask ourselves is why they did not feel that they needed to do so. The answer lies not in their beliefs about research *methods* but in their beliefs about the *purpose* of research, beliefs which are widely shared in information science. They say (p. 306),

> This study was designed to test the basic assumption of trigram analysis to determine if there is sufficient difference between the trigram compositions of correct and misspelled words for the latter to be reliably detected.

One cannot quarrel with this statement as far as it goes. It is certainly useful to know whether a technique can be used at all or not. We should note, however, that the authors are not concerned with *explaining* anything. They do not ask *why* trigram analysis works or how well it works in comparison to other methods. All they want to show is that their "machine" will work when they start it up. Since they have no interest in explaining why their "machine" works, they have no need to rule out "alternative explanations". For them, an experiment is not a controlled test of a hypothesis but only a demonstration of the workability of a technique. Zamora and his colleagues want to advance the state of a technology; they are not attempting to further the development of a science. Moreover, since the paper was accepted for publication, we can infer that the editors of the journal that published it agree that Zamora's goal was a proper one for a paper in information science.

We do not wish to criticize Zamora's study. It is a good one, given its goals. We also have no wish to quarrel with the standards of our discipline, but we would like to point out that research of this type has a very serious limitation: it can contribute nothing to the development of information science as a science. That is, it cannot lead to the production of any sort of theory. For, a scientific theory, as HOMANS[10] points out, consists of propositions about more or less general relationships among properties of nature. He says (p. 26),

> If we like, we can look on theory as a game. The winner is the man who can deduce the largest variety of empirical findings from the smallest number of general propositions, with the help of a variety of given conditions... But if theory is a game, it must like other games be played according to the rules, and the basic rules are that a player must state real propositions and make real deductions. Otherwise, no theory!

Researchers like Zamora and his colleagues are not concerned with stating or testing propositions about relationships between properties of nature, and so their research cannot contribute to the development of theory in information science. Anyone who has read much of the literature in this field knows that studies like that of Zamora *et al.* abound. Thus, it seems clear that one reason why true experimental studies are so rare in information science is that the prevailing standards of research in the discipline as mediated by the editors and reviewers of the field's prestigious journals do not require research to test theory. Since researchers are not much concerned with testing theories, they feel no need for experimental controls. However, this is not the only reason for the rarity of experimentation. Let us turn now to another.

Studies of phenomena not suitable for experimental manipulation

Many studies in information science deal with things that are so imbedded in their social settings that it is not easy to conceive how they might be abstracted to controlled, experimental situations. For instance, WILLIAMS[28] tests an interesting prediction about the "input of print and database products on database producer expenses and income". This is certainly an important issue for information scientists, but it is hard to imagine how it might be studied experimentally within the resources available to most researchers. Another example is Carpenter and Narin's[1] study of "The Adequacy of the *Science Citation Index* as an Indicator of International Scientific Activity". Their conclusion that the adequacy of the index varies among scientific fields and between countries with Roman and non-Roman writing systems is important and interesting. However, it is a historically specific phenomenon imbedded in its context and could not easily be studied in the experimental laboratory.

We should note, however, that there is an important difference between these two studies. Williams' study deals with a variable—the introduction of on-line bibliographic searching facilities—that is of a type that we earlier suggested was suitable for experimental study. It is inherently manipulatable, and we would like to know what it does in order to make the best use of it. Only we cannot study it experimentally because to do so would be too expensive.

In contrast, Carpenter and Narin's research deals with an altogether different sort of question. The current adequacy of the SCI as an indicator of scientific productivity is inherently a historically bounded question, and there is no reason to believe that in a controlled setting the results would be the same. To see why this is so, imagine what the term "current adequacy" would mean outside of a specific, historical situation. HAAS[7] gives a discussion of the problem of historical boundedness in social science.

The reason for drawing attention to this difference between the two studies is that research like that of Williams might be improved by the deliberate use of what CAMBELL and STANLEY[1] have called "quasi-experimental" designs which substitute other controls for those used by the experimenter when the latter are not feasible. We suggest that the adoption of quasi-experimental designs in cases where they are practical would greatly improve the theoretical usefulness of much research in information science.

Quasi-experimental research designs

Quasi-experimental designs are appropriate and useful when we want to make causal inferences from research in which experimental control, especially randomization, is not feasible. In such situations, we must make special efforts to rule out alternative explanations for our findings because we cannot be sure that our groups were equivalent to begin with. In thinking about how to rule out alternative explanations, it is well to remember that we do not need to rule out all conceivable alternatives but only those that are plausible ([13], p. 9). Thus, in designing a good quasi-experiment, we must first think through very carefully what plausible alternative explanations for our predicted findings there might be, and we must make special efforts to rule them out.

The basic strategy of the quasi-experiment is replication. If an effect can be observed to occur regularly in different groups, at different times, or under different conditions, then such alternative explanations as history, maturation or selection become less plausible. By the same token, if we can demonstrate repeatedly that in the absence of the 'treatment' the hypothesized effect does *not* occur, the attribution of causal efficacy to the treatment becomes more convincing.

One way of obtaining replication is through a *time series design.* This may be diagrammed as follows:

$$O \ O \ O \ OXO \ O \ O \ O.$$

In this design, repeated observations of the dependent variable are made at several times. Then, in the interval between two of the observations, the "treatment" is introduced, and then again repeated observations of the dependent variable are made. If we find a sudden, discontinuous change in O across the interval when the "treatment" occurred, we have a strong case for attributing causal importance to it. Note also that in this design it is not necessary that the "treatment" be introduced by the experimenter. It might be some naturally occurring event. For instance, the theoretical strength of Williams's conclusions concerning the effects of on-line bibliographic reference systems would be strengthened if she could show that the trends she observes are not simply continuations of preceding trends caused by something else entirely. This design eliminates most threats to the validity of a causal attribution except history. We cannot rule out the possibility that other events occurring during the time when the "treatment" occurred account for the observed changes. This is particularly important in a study like Williams's that took place over a considerable period of time.

A different approach is exemplified by the *nonequivalent control group design.* This design may be shown as follows:

$$\begin{array}{lll} O \ X \ O & \text{(Experimental group)} \\ O \quad\ \ O & \text{(control group)} \end{array}$$

This design is called the "nonequivalent control group" design because it does not include randomization to equate the experimental groups. Because of this lack, the persuasiveness of the design depends on the experimenter's success in convincing us on substantive (that is, nonstatistical) grounds that the two groups were in fact similar to begin with. This may be accomplished by the use of some sort of pretest (the first O).

The experimenter might also try to show directly that on "relevant" dimensions, the two groups were similar. For instance, in a study dealing with an information retrieval system, it might be shown that experimental subjects not receiving some special training were similar in educational background to those receiving the special training. This was done, for instance in MARCUS and REINTJES' [14] study of the on-line training component of their CONIT system. The difficulty is, of course, that it may not be possible to obtain groups that are truly equivalent in all "relevant" aspects or even to know which aspects are "relevant". Nevertheless, any sort of control group is better than none at all.

Moreover, this design has the advantage that it may be used with naturally occurring groups (e.g. school classes). Using such groups may often be less "reactive" or intrusive than creating special groups. Subjects in specially created groups are likely to have a strong awareness of being experimental subjects, and this feeling might cause them to behave differently from the way that they would have behaved outside of the experiment ([13], p. 13). In naturally occurring classes, on the other hand, the study might be integrated into the regular class work in such a way that the students need never know that a study was being performed. Of course, such a procedure raises ethical questions, but that is a separate issue.(See FILSTEAD [5] and DOUGLAS [3] for discussions of some of these ethical questions.)

Note that the nonequivalent control group design is strong where the time series design is weak, that is, in eliminating history as a plausible explanation for the observed effects. This suggests that an even stronger design might be created by combining the two. This is the goal of the *multiple time series design* which may be shown as follows:

$$\begin{array}{ll} O\ O\ O\ OXO\ O\ O\ O & \text{(experimental group)} \\ O\ O\ O\ O\ O\ O\ O\ O & \text{(control group)} \end{array}$$

This is a very strong design. It eliminates all threats to the internal validity of a causal attribution. It does not eliminate threats to external validity, but replications with different operationalizations of X or different measures of O would certainly make the attribution of external validity extremely strong.

We wish we had space to describe more of the quasi-experimental designs provided by Campbell and Stanley[1] and by Cook and Campbell[2]. Unfortunately, we lack the space to do so; we hope that our account has whetted your appetites enough that you will want to read their more detailed accounts. One word of caution is in order before closing. The statistical analysis of the results of the designs discussed here presents some special difficulties. The reader should therefore consult the relevant chapters in Cook and Campbell before undertaking a study using one of these designs.

REFERENCES

[1] DONALD T. CAMPBELL and JULIAN C. STANLEY, *Experimental and Quasi-Experimental Designs for Research*. Rand-McNally, Chicago (1963).

[2] THOMAS D. COOK and DONALD T. CAMPBELL, *Quasi-Experimentation*. Rand-McNally, Chicago (1979).

[3] JACK D. DOUGLAS, *Investigative Social Research*. Sage, Beverly Hills (1976).

[4] JEROME FELDMAN and WILLIAM R. SUTHERLAND, Rejuvenating experimental computer science. *Commun. ACM*. 1979, **22**, 497–502.

[5] WILLIAM J. FILSTEAD, (Ed.), *Qualitative Methodology: Firsthand Involvement with the Social World*. Markham, Chicago (1970).

[6] HAROLD GARFINKEL, *Studies in Ethnomethodology*. Prentice-Hall, Englewood Cliffs, New Jersey (1967).

[7] DAVID F. HAAS, Survey sampling and the logic of inference in sociology. *Am. Sociologist* 1982, **17**, 101–111.

[8] WILLIAM L. HAYES, *Statistics*. Holt, Rinehart & Winston, New York (1963).

[9] GEORGE CASPER HOMANS, Group factors in worker productivity. (Edited by HAROLD PROSHANSKY and BERNARD SEIDENBERG), *Basic Studies in Social Psychology*. Holt, Rinehart & Winston, pp. 592–604, New York (1965).

[10] GEORGE CASPER HOMANS, *The Nature of Social Science*. Harcourt, Brace & World, New York (1967).

[11] DAVID HUME, *An Inquiry Concerning Human Understanding*. (Edited by CHARLES W. HENDEL). The Liberal Arts Press, New York (1955).

[12] FRED N. KERLINGER, *Foundations of Behavioral Research*. Holt, Rinehart & Winston, New York (1964).

[13] EUGENE J. WEBB, DONALD T. CAMPBELL, RICHARD D. SCHWARTZ and LEE SECHREST, *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Rand-McNally, Chicago (1966).

APPENDIX: ARTICLES SURVEYED

(1) MARK P. CARPENTER and FRANCIS NARIN, The adequacy of the *Science Citation Index* as an indicator of international scientific activity. *JASIS* 1981, **32**, 430–439.

(2) CZESLAW DANILOWICZ and HENRYK SZARSKI, Selection of Scientific Journals Based on the Data Obtained From an Information Service System. *Inform. Proc. Management* 1981, **17**, 13–19.

(3) NANCY VAN HOUSE DEWATH, Fees for online bibliographic search services in publicly-supported libraries. *Library Res.* 1981, **3**, 29–45.

(4) CAROL HANSEN FENICHEL, Online searching: measures that discriminate among users with different types of experiences. *JANIS* 1981, **32**, 23–32.

(5) NANCY L. GELLER and JOHN S. DE CANI, Lifetime citation rates: a mathematical model to compare scientists work. *JANIS* 1981, **32**, 3–15.

(6) JAN DE GEUS, FRANS MULDER, BERT ZUURKE and MARILYN M. LEVINE, A replication of the Nelson and Mitroff experiment in teaching "bothsides" thinking. *JANIS* 1982, **33**, 76–81.

(7) LARRY HARDESTY, Use of library materials at a small liberal arts college. *Library Res.* 1981, **3**, 261–282.

(8) KERRY A. JOHNSON and MARILY DOMAS WHITE, The field dependence/field independence of information professional students. *Library Res.* 1981, **3**, 355–369.

(9) PAUL B. KANTOR, Levels of output related to cost of operation of scientific and technical libraries. *Library Res.* 1981, **3**, 1–28.

(10) CHAI KIM, Retrieval language of social sciences and natural sciences: a statistical investigation. *JASIS* 1982, **33**, 4–7.

(11) MANFRED KOCHEN, VICTORIA REICH and LEE COHEN, Influence of online bibliographic services on student behavior. *JASIS* 1981, **32**, 412–420.

(12) SUSAN O. LUNDBERG, A Delphi study of public library goals, innovations, and performance measurements. *Library Res.* 1981, **3**, 67–90.

(13) TAKASHI MAEDA, An approach toward functional text structure analysis of scientific and technical documents. *Inform. Proc. Management* 1981, **17**, 329–339.

(14) RICHARD S. MARCUS and J. FRANCIS REINTJES, A translating computer interface for end-user operation of heterogeneous retrieval systems. II. Evaluation. *JANIS* 1981, **32**, 304–317.

(15) PENELOPE MCKEE, Weeding the Forest Hill branch of Torono public library by the Slote method: a test case. *Library Res.* 1981, **3**, 283–301.

(16) ELLIOT NOMA, Untangling citation networks. *Inform. Proc. Management* 1982, **18**, 43–53.

(17) MIRANDA LEE PAO, Collaboration in computational musicology. *JANIS* 1982, **33**, 38–43.

(18) MIRANDA LEE PAO, Co-authorship as communication measure. *Library Res.* 1980, **2**, 327–338.

(19) BLUMA C. PERITZ, Citation characteristics in library science: results from a bibliometric survey. *Library Res.* 1981, **3**, 47–65.

(20) MICHAEL ROGERS and MARY E. SAPP, Selected roles of information goods and services in the U.S. National economy. *Inform. Proc. Management* 1981, **17**, 195–213.

(21) MICHAEL ROGERS and ELIZABETH TAYLOR, The U.S. information sector and GNP. *Inform. Proc. Management* 1981, **17**, 163–194.

(22) WILLIAM B. ROUSE, SANDRA H. ROUSE and DAVID R. MOREHEAD, Human information seeking: online searching of bibliographic networks. *Inform. Proc. Management* 1982, **18**, 141–149.

(23) MARY E. ROWBOTTOM and PETER WILLETT, The effect of subject matter on the automatic indexing of full text. *JASIS* 1982, **33**, 139–141.

(24) HENRY SMALL, The relationship of information science to the social sciences: a cogitation analysis. *Inform. Proc. Management* 1981, **17**, 39–50.

(25) MICHAEL V. SULLIVAN, BETTY VADEBONCOEUR, NANCY SHIOTANI and PETER STANGL, Obsolescence in biomedical journals: not an artifact of literature growth. *Library Res.* 1981, **2**, 29–46.

(26) HOWARD D. WHITE, Cocited author retrieval online: an experiment with the indicators literature. *JASIS* 1981, **32**, 16–21.

(27) WAYNE A. WIEGAND and GERI GREENWAY, A comparative analysis of the socioeconomic and professional characteristics of American Library Association Executive Board and Council Members, 1876–1917. *Library Res.* 1980, **2**, 309–325.

(28) MARTHA E. WILLIAMS, Relative impact of print and database products on database producer expenses and income—trends for database financial analysis. *Inform. Proc. Management* 1981, **17**, 263–276.

(29) ROBERT V. WILLIAMS, Sources of the variability in level of public library development in the United States: a comparative analysis. *Library Res.* 1980, **2**, 157–176.

(30) E. J. YANNAKOUDAKIS, P. GOYAL and J. A. HUGGILL, The generation and use of text fragments for data comprehension. *Inform. Proc. Management* 1982, **18**, 15–21.

(31) E. M. ZAMORA, J. J. POLLOCK and ANTONIO ZAMORA, The use of trigram analysis for spelling error detection. *Inform. Proc. Management* 1981, **17**, 305–316.