The Use of Computer-Monitored Data in Information Science and Communication Research

Ronald E. Rice
Annenberg School of Communications, University of Southern California, Los Angeles, CA 90089

Christine L. Borgman
Institute for Communication Research, Stanford University, Stanford, CA 94305

As computer-based information retrieval and communication systems become more commonplace, researchers have a greater opportunity to evaluate the uses and impacts of new communication technologies. The systems and the kinds of data now available are discussed along with advantages and disadvantages of using computer-monitored data.

Introduction

Computerized systems are becoming an integral part of our daily lives. We use them to support routine tasks, from text processing to records management. But now we are beginning to use them to support a most fundamental task: human communication. Computer systems and their networks are new media for human interaction. Furthermore, they can also provide information about how a medium is used and what it transmits. It is appropriate, then, to consider the prospects and problems presented by the ability of computer-based systems to monitor or collect kinds of data as humans use such systems to exchange and retrieve information. It is our premise that computer-monitored data provide new opportunities and responsibilities for researchers.

Computer monitoring of an information or communication system is simply the automatic logging of the type, content, or time of transactions made by a person from a terminal with that system.

Penniman and Dominick write, "Monitor data can provide: 1) direct and immediate diagnostic aid to the user; 2) grouped or individual user evaluation, e.g., analysis of user performance, user success/satisfaction; 3) system protection, e.g., diagnosis of attempts at unauthorized system access; and 4) system evaluation" [1, p. 23]. Concentrating on the first two categories,* it seems timely to provide a review and overview of the following issues of computer-monitored data for research in communication and information science: (1) Description of systems and kinds of data possible; (2) Uses of the data; (3) Advantages of computer-monitored data; (4) Problems with computer-monitored data and collection; (5) Prior research using computer-monitored data; and (6) Prospects for future research.

Description of Systems and Kinds of Data Possible

Most of the monitoring studies of interest have involved either information retrieval systems or computer-based communication systems.

Information Retrieval Systems

An online information retrieval (IR) system is "one in which a user can, via computer, directly interrogate a machine-readable database of documents or document representations" [10, p. 1]. Information retrieval systems can be classified either by the type of database on which it operates or by search capabilities [11,12,7,13]. Bibliographic databases include descriptions and accessibility of literature, such as journal articles or books. Nonbiblio-

Received August 27, 1982; accepted October 28, 1982

*Readers interested in the other categories may wish to consider systems performance or cross systems interfacing literature and the telecommunications traffic, cycle efficiency, and load management literature [2,3]. Or, they may be interested in human-computer interface issues in computer simulations or human factors experiments [4-8].

graphic databases include everything else, such as statistical files, company records, stock exchange activity, etc.

The most sophisticated types of information retrieval system allow searching on a large number of data elements combined by Boolean logic relationships. Other capabilities such as truncating terms and searching on word stems are frequently available. Commercial retrieval systems such as DIALOG, SDC ORBIT, and Lexis fall into this category.

Less sophisticated are the search key systems which require an exact match on a few characters of input. For example, an author search might be restricted to searching on a key of the first five characters of the author's surname plus the first four characters of the first name. Most "online catalogs" which are used in place of card catalogs use this structure. These systems are much less flexible than the Boolean logic systems and have fewer searchable data elements and fewer commands.

More recently, a class of IR systems aimed at the consumer market has developed. This class includes videotex and teletext systems which typically require little or no training for use. (See refs. 14–16 for introductions to the technology and applications.) Videotex is a generic term for systems "that transmit text and graphics to the business or home viewer by means of signals carried over a telephone line, cable, or any of the TV or radio broadcast channels" [17].

Many public, academic, and private libraries are replacing their card catalogs with online catalogs, which are also IR systems [18,19]. The online catalogs are bibliographic retrieval systems similar to those described above, but are designed for use by library patrons. Most of the systems are much simpler to use than the multiple database retrieval systems such as DIALOG. Several other monitoring studies online catalogs are now in progress [20-22]. Many IR systems exist in private organizations and government for the purpose of maintaining internal records and various kinds of management information. Monitoring studies of these systems are rarely done for the purpose of studying user behavior.

Computer-Based Communication Systems

A computer-based communication (CBC) system is a generic term for electronic messaging or computer conferencing as well as for functions of more sophisticated knowledge worker augmentation systems. Using a CBC system, people can send text (data, memoes, letters, reports, etc.) to one another (to one or several individuals or predefined groups of individuals) or may share files (conference comments, working drafts, co-authored papers, programs, etc.) while communicating (usually not at the same time) via geographically dispersed terminals (video or hard-copy) using a shared host computer and telecommunications lines (national networks, cable, etc.) [23-32]. Computer conferencing systems can *structure* the group communication patterns according to the needs of the users. Another extension of CBC systems is the "electronic journal" [33].

Types of Data Possible

Online monitoring produces large datasets. The precise data elements collected will vary by the type of system, but most data are either transactional—what happened when, where, and by whom—or temporal—what was the pacing of the interaction, how much time is spent in certain activities, and so on. Some of the categories of computer-monitored data which seem to promise insights, along with comments as to their meaning and relevance for research, follow.

By collecting a few well-defined "primitive" measures we can generate a variety of useful aggregate and ratio measures of communication behavior. At a minimum, we can collect most of the following data points from both IR and CBC systems:

- terminal and user identification numbers (ID can be unique or can represent classes of system users, such as organizational division, hierarchical level, external groups, etc.)
- user start time, end time
- protocol commands (specific to interaction languages) (send message, perform author search, examine user directory, print, store, transfer data/message)
- full text content (of search, result, message, etc.)
- topic, keyword as content summary
- code or response by system
- messages received or waiting: number, length, type
- message audience category (single or multiple, private or public, draft workspace or conference, etc.)
- number of matches retrieved (including none)
- errors (bad entries, illegal commands, no target record)
- user entries in response to computer-administered queries

Many other portions of system transactions can be captured, depending on system resources and researcher interest. Some comments on the choice of data elements, with respect to three kinds of analysis—*pattern*, *error*, and *time* analysis—are in order.

We can use computer-monitored data to uncover and describe *patterns* of system usage. It is important to distinguish between terminal identification number and user identification number, as multiple users may share a terminal or the same user may use different terminals, possibly blurring pattern distinctions across users. Computer-monitored data from IR systems are usually organized into individual user sessions—from the time a user gives the first command, through the last command given. Each command is treated as a transaction and each response from the computer may be treated as another transaction; consecutive sets of user commands and system responses are normally treated as pairs for analysis purposes. Generally, a string of such commands can be aggregated to comprise a search task, while all such commands and results within a single user session may be aggregated to constitute a "session" datum. The data are usually used to look at user behavior within individual sessions and to look for comparisons between users, perhaps across systems or tasks. The session approach allows the study of individual user behavior, whereas aggregating all the transactions from a system without discriminating between user session allows the study of system use. In CBC systems one research goal may be to describe or model patterns of communication networks, built up from links between senders and receivers, as described below.

We can focus on the analysis of the frequency, type, and context of *errors*. We make a distinction between the actual protocol command given, which can be recorded as a code from a standard list of commands, and the actual text of the command (e.g., what was searched for, text of message sent). However, a particular system can identify only certain types of errors (incorrect entries, commands out of sequence, etc.). The system will typically miss logical errors and some content errors which appear correct (e.g., message was sent to wrong person; wrong item was searched for, but matches were made).

We can focus on issues of *timing* and duration. Some caution must be taken in interpreting user start and end times. The system record of start and end time will show when the command reached the system, based on the system clock time. The system clock may not be the same as the real clock (an important consideration if any observers are collecting data concurrently). Depending on the system's response time, the recorded time may be later than the time at which the command was actually issued. The elapsed time between transactions may provide useful information about an individual's "think time" associated with specific commands and identify some of the more difficult-to-use features in the system. Several difficulties arise with this approach, however. For example, the "think time" associated with the first command (prior to starting the search) and the last command (after it has been issued) cannot be captured explicitly. Therefore, the researcher must attribute elapsed time consistently to the preceding or following command, or find some way to split it between the two. Also, the choice of time interval affects the meaning of the analysis unit. For example, how quickly does a message recipient have to respond for the message to be considered "answered" or "reciprocated"?

Once the basic data elements have been parsed into user sessions, some aggregate measures can be computed for use in pattern, error, and time analysis. Examples from each analysis include the frequency of individual commands or the equality of participation in group communication; the ratio of successful to total commands or the level of group consensus; and the elapsed time between commands or the time for a group to reach a decision. A number of analytical techniques may then be appropriate [1,34,35]. One approach is to analyze the data in terms of strings or patterns of commands using a Markov process approach. This allows the identification (but not necessarily the explanation) of patterns which lead to successful and unsuccessful searches or transitions from one communication role to another, and may identify the location and context of errors. The combination of pattern analysis, error analysis, and time analysis can lead to a helpful model of IR/CBC user behavior. The results can be used to inform system improvements, to make online suggestions if the system suspects that the user is having problems, or to compare with similar offline behavior.

Uses of the Data

Information scientists and communication researchers will be concerned with two analytical purposes for such data:

(1) *evaluations* of the uses and utility of such communication media by individuals (in their social settings, such as organizations or research groups), and
(2) indicators of the *impacts* of such communication media upon the users and their social settings.

Monitoring for Evaluative Purposes

With respect to evaluations of IR and CBC systems, there have been several major reviews of the issues. Reviewing the applied systems design literature, Rouse [36] integrates research results from multiple disciplines (such as cognitive psychology, physiology, social psychology, engineering and the like) to suggest ways in which system evaluations can improve man–computer interaction. Moran [37] takes a more theoretical, cognitive psychology perspective and proposes the field of "user psychology"; and later [38] proposes a "command language grammar" for the user interface and suggests a psychological user model for viewing the user. Online monitoring is one empirical method for building and testing such models. The complementary review by Paisley of what is actually involved in information work is based upon his belief that, "In order to think about improving the system of information transactions related to work, it helps to develop a framework for specifying *who* is processing *what kinds* of information for *what purposes*" [391, p. 159].

Computer-monitored data are clearly only one source of information for system evaluators and researchers, and implications from such data are only one component of a thorough evaluation. Any evaluation of such systems consists of some small intersection of dimensions including the *stakeholders*, *evaluation goals or criteria*, and *analysis domain*. "Once the evaluator understands which cells [from this planning matrix] are being considered in the evaluation effort, some approaches and methods [as well as types of data] become quite appropriate and even elegantly suited to the topic at hand" [35]. Other issues include historical and methodological approaches [1,40,41], organizational objectives [2], and multiple evaluation perspectives [34,42].

Monitoring for Impact Analysis

Assessing IR and CBC impacts is another purpose of using computer-/monitored data. Empirical impacts of computer-mediated interpersonal and organizational communication [29,43] or the theoretical foundations of such research [44] can be used to help avoid negative impacts and to alter prejudicial attitudes which potential users may have about such systems. In this sense, evaluation research and impact research are interconnected: Prior knowledge of potential impacts and typical usage patterns establishes baselines for later comparisons and for initial system design and organizational planning. As Bair [45] writes, "several years of studying the impact of office automation on productivity have shown that changes in the office communication system are the most important for productivity improvement ... [subject to] many conditions in system and human and organization factors." Ongoing evaluation and impact research guides the development of the system (perhaps even to enable the system to "learn" about the user and offer appropriate interfaces) as well as reveals the kinds of behaviors and attitudes users may have toward the system.

Monitor data can be used either for predictive or description impact analyses. The researcher can begin with a model or theoretical framework and test specific hypotheses, or can take an exploratory approach and attempt to describe the behavior of people using new media.

Some Evaluation and Impact Questions

We note here some questions in IR and CBC research: How much time is spent in different activities, such as selecting terms, forming search strategies, reviewing output, composing and responding to messages, or arriving at group decisions? What factors discriminate between patterns of user behavior (such as early system acceptance, heavy use, complexity of command use, etc.)? What is a model of effective (or ineffective) search behavior? What models of group interaction apply to communication via a CBC system? How does online behavior compare to comparable offline behavior (such as searching an online versus a card catalog, or CBC versus face-to-face communication)? How does online behavior vary between systems or different communication structures or group tasks?

Advantages of Computer-Monitored Data

Automated Collection

From a logistical point of view, perhaps one of the greatest advantages of collecting and using computer-monitored data is that the computer does the collecting! This is no small task when one considers how difficult it is to collect detailed data from many subjects over time. More to the point, the researcher is unlikely ever to be able to collect such data. In addition, having accessible, computer-collected and -maintained large, complex databases encourages reanalysis by other researchers with differing perspectives. Replications and meta-evaluations are then more likely.

Unobtrusive Collection of Accurate Data

Unlike many questionnaires, field experiments or controlled experiments, the collection of computer-monitored data typically involves littler or no response bias or demand characteristics from the subjects. It is essentially unobtrusive, which may increase the validity of the data [46]. Experiments run on a system are replicable, the timing of commands may be controlled, questions can be randomized, etc. Consider also the wealth of recent research which indicates that, in general, respondents' *reports* of their communication activities diverge widely from their *actual* communication behavior as observed or monitored [47–50]. Although the controversy continues as to the extent, form, and generalizability of these kinds of divergence, one recent study is very convincing [51]. An experiment administered through the main computer of a computer conferencing system asked users typical sociometric questions, and compared these data to the communication behavior monitored by the computer. The results not only reaffirmed the typical presence of large discrepancies between reported and actual communication behavior (both in forgetting and inventing recipients of messages), but also showed this discrepancy even in reports collected within minutes of a respondent's sending the message.

Full Census and Network Data

Another, related aspect of computer-monitored data, is that these accurate census data allow us to investigate the communication networks of groups of users. Rogers and Kincaid [52, p. 346) define communication networks as consisting of *relations* among "... interconnected individuals who are linked by patterned flows of information." These networks link organizations and user groups with each other and with the environment and are, in fact, one picture of an organization's or group's structure [53–58]. Rice and Danowski [35, p. 324] write:

> A network-oriented evaluation would measure communication flows ... before, during, and after the

implementation of a new electronic messaging system, for example, to determine whether the technology assists the development of desired communication flows, whether other organizational media ... are affected, whether certain tasks are performed better in these altered communication patterns, whether the same information can be handled in fewer transformations among media ... whether decision-making is centralized or decentralized...

Longitudinal Data

The servicing computer can capture extensive *longitudinal* network data so that researchers avoid the ungrounded assumption of much cross-sectional research that the system under study is at some equilibrium state. These data may be discrete or continuous; the notion of analyzing continuous time-dependent communication processes is perhaps foreign to most communication researchers precisely because obtaining such data is so difficult [59].

Automated Experiments

Finally, the same computer which provides the facilities for human information exchange or retrieval can also administer controlled experiments, collect the data directly, "document the problems, and the decisions made at each stage," and follow up the experiments with joint on-line authorship of reports (25). Some examples are reported in the references by Hiltz and her colleagues [60–66]. See also refs. [67] and [68] and issues of *Behavior Research Methods and Instrumentation.*

Problems with Computer-Monitored Data and Collection

In general, the advantages of computer-monitored data are unobtainable in scope (extent of population, length, detail) or in type (kinds of data from traditional data-collection procedures, due either to problems of accuracy or sheer inaccessibility. The same attributes of these data which can be called a researcher's boon, however, may also be a researcher's bane.

Extensive Datasets May Mean Extensive Data Management

The very fact that massive amounts of particulate data can be collected means that someone has to manage all those data! This has serious implications in terms of *budgets, time,* and *expertise. Budgets,* because preprocessing these data may take quite large sums of computer time, both in the on-going day-to-day collection and in the conversion of the raw data into analyzable datasets. Evaluators need to integrate plans for using such data into system design to minimize later processing requirements.

Time, because as anyone who has had to handle computer tapes, multiple datafiles, and custome-developed programming knows, these complex operations generally mushroom into time-consuming activities. Much raw data can only be transformed into analyzable units through preanalysis. For example, online bibliographic searching sessions on public terminals may have no clear demarcation between user sessions, or the stream of raw search data must be parsed into command strings and user sessions based upon knowledge gained from related data [20,21]. Even when the computer does some preprocessing it may impose arbitrary aggregation of data.

Expertise, because some member of the research team must know how to program or execute the necessary routines and transfer large sets of data. For example, in creating who-to-whom network matrices from messaging data or transition matrices of commands from on-line searching, it is useful to know that core storage limits can be reached quite easily using full word-lengths, but that orders of magnitude storage requirements can be saved using two or four bits to represent categories of data. In addition, it happens frequently enough that a row-by-column data structure (the usual case-by-variables form of data) has to be transformed into a column-by-row structure for certain programs (such as time-series, or some network analysis programs). This maneuver can be a laborious and error-prone data entry task or even a modest programming job. However, one can easily use SAS' MATRIX procedure to transpose the initial matrix and write out the new one. Penniman and Dominick [1, p. 23] strongly recommend that researchers "store monitor data in [database management systems] capable of interfacing with external software."

Another aspect of expertise is that the researcher needs perhaps to be more systematic and theory/hypothesis-driven than usual. Sorting the study down to a manageable set of questions to pursue is a much greater problem in a monitoring study than is a lack of data.

Ethics and Privacy

One computer impact salient to the average person is the storage and use of data on that person [69,70]. Only in a few instances is it appropriate to collect full text records. People have the right to exchange information in privacy. In the information society where information is power and contacts constitute major parts of our social structure, there are some who might claim that *who* they communicate with should remain even more private. In other cases it is clear that textual content is *meant* to be public, and analysis is less questionable [19,71,72]. Most extant computer-monitored data extant come from either pilot research (which is usually publically funded) or from within organizations which wish to utilize their systems more effectively. In both cases, the users and the implementors have a keen interest in understanding system use and impacts. However, we emphasize that subjects must have the right to deny permission for access to portions of the data, although such datasets collected under auspices of public funding must be made publicly available. Individuals' identification numbers may be randomized before analysis or before release to maintain confidentiality [73,74]. A project may limit data collection to the users' commands, with content bypassed. For example, an analysis plan could retain the information that a user made a title search in a bibliographic system, but not the search key or any data on what was searched for.

Accurate, But of What?

Computer-monitored data obviously do not portray the whole picture of human communication. Studies clearly show the social power and utility of very informal, unmonitorable organizational communication [75]. We know that much of our human communication occurs on the nonverbal level. And, as the literature on respondent inaccuracy notes, perhaps people do not report their behavior accurately, but they do apparently base their decisions upon their *attitudes* anyway. All this is true. But if we are to understand the actual use of such systems, and the impacts of exchanging specific kinds of information, we must study behavior at least as much as attitudes.

Prior Research Using Computer-Monitored Data

Information Retrieval Systems

Information retrieval monitoring studies have a considerable history (particularly in performance evaluation, which we do not address). Penniman [76] and Penniman and Dominick [1] provide a thorough review of IR monitoring studies through 1980. While they do not explicitly list results for these studies, they do outline some of the variables which were studied: interactions per session, interactions per minute, probability of session length in minutes for different database types [77]; query complexity [78]; learning curve for terminal users [79]; and most likely paths of next actions in a search session [80,81].

We are beginning to develop an understanding of user behavior—at least of *differences* among users [66]. In his dissertation, Penniman described users of the BASIS system in terms of Markov matrices. He found that the user cycles through a "semantic mode of establishing a common vocabulary during the first portion of the session" [81, p. 152]. After about four interactions, the typical user concentrates on syntax and formulates a logical search strategy. The user may then either return to the vocabulary mode, or else displace documents to test the search hypothesis. Beyond this point, user behavior cycles in a changing pattern which can be approximated by the Markov model. In a more recent study [73], Penniman compared patterns of use for infrequent, moderate, and frequent users of the MEDLINE system. He found significant differences between groups on frequency of individual commands. Infrequent users were slower; experienced searchers used more (and the more advanced) commands and more connect time. Chapman [82] compared users trained by a human to users trained by a computer intermediary, finding that inexperienced and novice users tend to follow the suggestions of the teacher, whether human or computer. Amount of searching experience had the greatest effect on the user's information seeking behavior, however.

Insights from actual system usage can have important policy implications. In the one publicly-funded US videotex evaluation funded so far, the University of Kentucky/USDA Green Thumb system [83,84] provided system data which correlated less than 30% with reported usage figures while showing a clear downward trend over the 13 months recorded. Figure 1 shows the downward trend in system usage in the two pilot counties involved. USDA policy and marketing implications based upon system-monitored system data ran counter to some of the implications based upon respondent's reports of their usage. Several monitoring studies on online catalogs

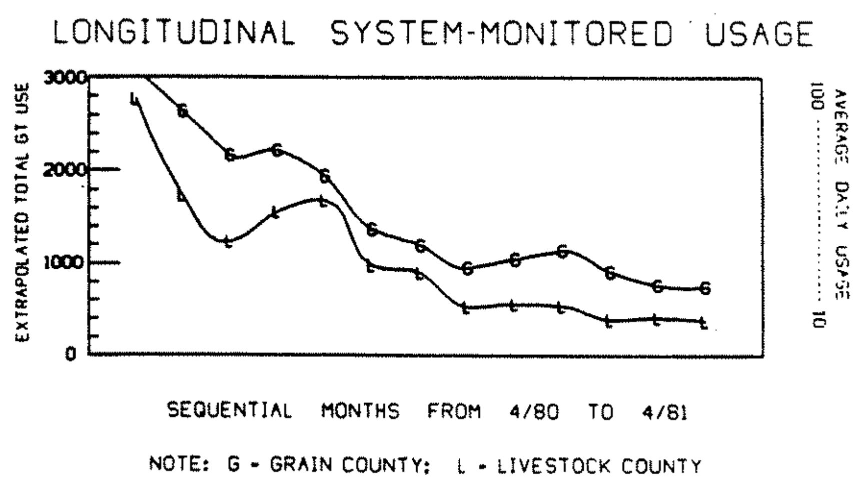

FIG. 1. Monthly green thumb usage.

have the express intention of aiding library policy decisions [20-22,85], while similar systems are evaluated for medical research purposes [86].

Electronic Messaging and Computer Conferencing

Several studies have used computer-monitored communication data to evaluate the utility and impacts of electronic messaging systems: See ref. [31], [87-102], and others referenced elsewhere [43].

Some studies have revealed the organizational communication functions which employees perform via electronic messages, or the progress of developmentally disabled children as they communicate via personal computer. Others have revealed tendencies in the direction of communication flows (usually increased flows, with a tendency for more downward or horizontal flows). A few have shown that substitution for other media, usage, and reported benefits depend on the nature of the organizational units (such as managerial versus sales representatives). Still others have replicated early studies which showed little relationship between initial attitudes (except perhaps for users' motivation to use the system, expected later usage, and prior knowledge of the status of other users) and later actual usage. Network analyses of usage data or network measures included with reported data have revealed that the same users tend to increase the number both of others to whom they send and of others from whom they receive messages, and that past system usage is (not surprisingly) related to future system use.

There is also a growing amount of research using computer-monitored data from computer conferencing systems: See refs. [19, [23], [25], [29], [40], [61-63], [65], [71], [72], [91], [103-108], and others noted elsewhere [74].

Under the aegis of Hiltz and Johansen there are now massive summaries of the vast experience to date on the acceptance, use, and impacts of such systems [25,65, 104]. Computer-monitored usage (and sometimes computer-administered questionnaire) data were instrumental in detecting the evolution of communication behavior and attitudes; the necessity of a leader in some kinds of tasks for successful use; comparisons of the effects of media on task-solution, consensus and quality of decision; the role of motivation and access to peers in the success and impacts of such systems; the effects of idea-generation and productivity in research communities; etc.

One analysis of 24 months of data from a nationwide computer conferencing system showed how communication network indices could provide indications of potential group effectiveness and cohesiveness. Figure 2 [74] shows the trend over time in the ratio of within-group to total message links for the group comprising system programmers and user consultants. Over time, this group fulfilled its task as a "service" group, moving from communicating within itself to exchanging information with external system users.

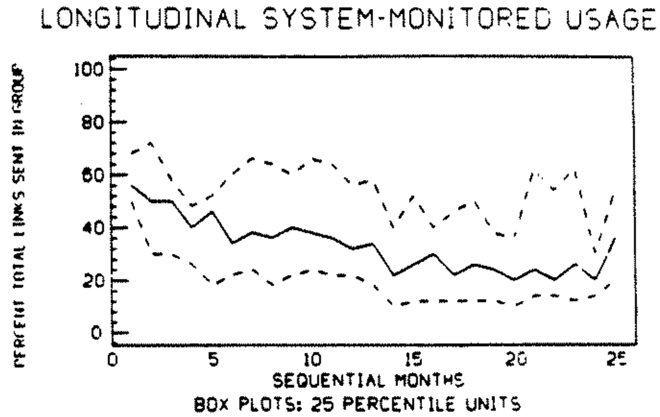

FIG. 2. Ratio of within-group links sent to total links sent, EIES group 10 and 99.

Infomedia [105] provides a commercial computer conferencing service which can analyze computer-monitored data to advise client organizations on the use of the system, to detect patterns of information exchange associated with group stability, success and survival, and to redesign the system itself.

Prospects for Future Research

Theoretical perspectives and research agendas on unmediated information retrieval and human communication may be profitably applied to user behavior in IR and CBC systems. Information retrieval is a fundamental component of performing tasks and making decisions. Paisley's [39] conceptual review of what is involved in information work suggests evaluative typologies and testable processes for which computer-monitored data might be useful. For example, the *functional kinds* of information used in executing a plan for accomplishing a task may be automatically measured, either by content analysis or counts of keywords and indices.

Bales [109] showed that group interaction proceeds in stages. An important impact question is whether, and under what conditions, groups engage in similar processes when using CBC systems. Hiltz and her colleagues have begun to answer this question, finding similarities in group interactions in face-to-face and CBC environments, but specifying different conditions and results (such as higher correctness of decision but less overt consensus when the group task is to solve a problem). A related area of communication research—conversational sequencing—has already established a tradition of using computer-monitored data. Cappella and Streibel [110] use a computerized system to collect and digitize the audio portion of conversations to test theories on how individuals establish and change their timing, duration, and pattern of verbal interaction.

Another theoretical framework with implications for interaction via CBC systems is Blau's [111] notion of inequality and heterogeneity in social structure. Some of his earlier work dealt specifically with bureaucratic and organizational social structures, exactly the area likely to be affected soonest by CBC systems. A tentative test of some of these notions found that reciprocal relations are an important attribute of a CBC system's social structure [107]. Looking forward to extensive use of CBC systems in offices, Ellis and Nutt [24] note current research problems in office information systems; some of these could be illuminated by computer-monitored data.

We must also understand the reasons for the discrepancy between user reports of system usage and actual usage behavior, but particularly specify how they are related in the context of mediated systems. When communication partners can retrieve others' past (forgotton or disputed) daily discourses from the computer, what will become of notions of trust and reliance upon another's "word"? If designers plan or managers operate systems based upon actual computer-monitored usage, but users evaluate systems based upon perceived usage, who should have the final decisions? Will communication research have to admit to two complementary versions of communication behavior?

From a technical perspective, it seems clear that computer-monitored data from IR and CBC systems are useful in evaluation and impact research, but that the actual extraction and management of such data may be difficult. Therefore, we would hope that designers, managers, and researchers devise ways to integrate data-collection and summary reporting into IR and CBC systems as a natural function of system operation. In addition, institutions which grant funding for system evaluations and research might wish to consider including monies for the creation and documentation of clean datasets for later public use. This service is as important as the more traditional diffusion of research knowledge by means of journal articles and the like.

References

1. Penniman, W.; Dominick, W. "Monitoring and evaluation of on-line information system usage." *Information Processing and Management*. 116:17–35; 1980.
2. Hamilton, S.; Chervany, N. "Evaluating information systems effectiveness—Part 1: Comparing evaluation approaches." *MIS Quarterly*. 5(3):55–70; 1981.
3. Svoboda, L. *Computer Performance Measurement and Evaluation Methods*. New York: Elsevier; 1976.
4. Card, S.; Moran, T.; Newell, A. *The Psychology of Human-Computer Interaction*. New York: Erlbaum; 1982.
5. Embley, D.; Nagy, G. "Behavioral aspects of text editors." *Computing Surveys*. 13(1):15–70; 1981.
6. Ramsey, H.; Atwood, M. *Human Factors in Computer Systems: A Review of the Literature*. Englewood, CO: Office of Naval Research, Technical Report SAI 79 111 DEN; 1979.
7. Reisner, P. "Human factors studies of database query languages: A survey and assessment." *Computing Surveys*. 13(1): 13-31; 1981.
8. Rouse, W. "Human-computer interaction in the control of dynamic systems." *Computing Surveys*. 13(1):71–99; 1981.
9. Shneiderman, B. *Software Psychology*. Cambridge, MA: Winthrop; 1980.
10. Lancaster, F. W.; Fayen, E. G. *Information Retrieval On-Line*. Los Angeles: Melville Pub. Co.; 1973.
11. Borgman, C. L.; Moghdam, D.; Corbett, P. K. *Effective Online Searching: A Basic Text*. New York: Marcel Dekker, in press.
12. Meadow, C. T.; Cochrane, P. A. *Basics of Online Searching*. New York: Wiley; 1981.
13. Williams, M. E.; Lannom, L.; Robins, C. G. *Computer-Readable Databases: A Directory and Data Sourcebook*. White Plains, NY: Knowledge Industry Publications; 1982.
14. Cherry, S. "Telereference: The new TV information systems." *American Libraries*. 11(2):94–110; 1980.
15. Pye, R. "Information retrieval services." *Telecommunications Policy*. 2(1):73–78; 1978.
16. Woolfe, R. *Videotex: The New Television-Telephone Information Services*. Philadelphia: Heyden and Son; 1980.
17. Bloom, L. R.; Hanson, A. G.; Linfield, R. F; Wortendyke, D. R. *Videotex Systems and Services*. Boulder, CO: US Dept. of Commerce, National Telecommunications and Information Administration, Institute for Communication Sciences, NTIA Report 80-50; October 1980.
18. Hildreth, C. R. *Online Public Access Catalogs: The User Interface*. Dublin, OH: Online Computer Library Center; 1982.
19. Ferguson, J. "PLANET: A computer conferencing system and its evaluation through a case study." *Behavior Research Methods and Instrumentation*. 9(2)92–95; 1977.
20. Borgman, C. L. Work in progress. Stanford, CA: Stanford University; 1982.
21. Tolle, J. Work in progress at OCLC, Inc., Dublin, OH; 1982.
22. Lawrence, G. S. Work in progress at University of California (Berkeley), Division of Library Automation; 1982.
23. Bamford, H. "Computer conferencing: The exchange of experience." *Telecommunications Policy*. 4(3):215–220; 1980.
24. Ellis, C.; Nutt, G. "Office information systems and computer science." *Computing Surveys*. 12(1):27–60; 1980.
25. Hiltz, S. R.; Turoff, M. *The Network Nation: Human Communication Via Computer*. Reading, MA: Addison-Wesley; 1978.
26. Newell, A.; Sproull, R. "Computer networks: Prospects for scientists." *Science*. 215(4534):843–851; 1982.
27. Panko, R. "The EMS revolution: a survey of electronic message systems." *Computerworld*. 14(34):45–56; 1980.
28. Panko, R. "Standards for electronic message systems." *Telecommunications Policy*. 5(3):181-197; 1981.
29. Rice, R. E. "The impacts of computer-mediated organizational and interpersonal communication." In: M. Williams, Ed. *Annual Review of Information Science and Technology, Volume 15*. White Plains, NY: Knowledge Industry Publications; 1980b: 221-249.
30. Uhlig, R. (Ed.) *Proceedings of the IFIP TC-6 International Symposium on Computer Message Systems*. New York: Elsevier North-Holland; 1981.
31. Uhlig, R.; Farber, D.; Bair, J. *The Office of the Future: Communications and Computers*. New York: North-Holland; 1979.
32. Zinn, K. "Computer facilitation of communication within professional communities." *Behavior Research Methods and Instrumentation*. 9(2):96–107; 1977.
33. Turoff, M.; Hiltz, S. R. "The Electronic Journal: A Progress Report." *Journal of the American Society for Information Science*. 33(4):195–202; 1982.
34. Dominick, W. "User interaction monitoring as a multi-level process." *Proceedings of Society for Information Science*. 14:63; 1977.
35. Rice, R. E.; Danowski, J. "Issues in computer conferencing evaluation and research." In: S. R. Hiltz; E. Kerr, Eds. *Studies of Computer-Mediated Communications Systems: A Synthesis of the Findings*, 1981; 294-326.
36. Rouse, W. "Design of man–computer interfaces for on-line interactive systems." *Proceedings of the IEEE*. 63(6):847–857; 1975.
37. Moran, T. P. "An applied psychology of the user." (Guest Edi-

tor's Introduction, Special Issue on the Psychology of Human-Computer Interaction.) *ACM Computing Surveys.* 13(1):1-12; 1981a.

38. Moran, T. P. "The command language grammar: A representation for the user interface of interactive computer systems." *International Journal of Man-Machine Studies.* 15(1):3-50; 1981b.
39. Paisley, W. "Information and work." In: B. Dervin; M. Voigt, Eds. *Progress in Communication Sciences, Vol. II.* Norwood, NJ: Ablex; 1980:114-165.
40. Johansen, R.; Miller, R.; Vallee, J. "Group communication through electronic media: Fundamental choices and social effects." In: H. Linstone; M. Turoff, Eds. *The Delphi Method: Techniques and Applications.* Reading, MA: Addison-Wesley; 1975:577-534.
41. Urban, J.; Dominick, W. "Design and implementation considerations for monitoring and evaluating information systems." *Proceedings of the Southeastern Regional ACM Conference.* 356-370; 1977.
42. Carlson, E. "Evaluating the impact of information systems." *Management Informatics.* 3:56-67; 1974.
43. Rice, R. E. "Computer conferencing." In: B. Dervin; M. Voigt, Eds. *Progress in Communication Sciences, Volume 2.* New York: Ablex; 1980a:215-240.
44. Kling, R. "Social analyses of computing: Theoretical perspectives in recent empirical research." *Computing Surveys.* 12(1):61-110; 1980.
45. Bair, J. "Communication in the office of the future: Where the real payoff may be." *Business Communications Review.* 9(1):1-11; 1979.
46. Webb, E.; Campbell, D.; Schwartz, R.; Sechrest, L. *Unobtrusive Measures: Nonreactive Research in the Social Sciences.* Chicago: Rand-McNally; 1966.
47. Berger, C.; Roloff, M. "Social cognition, self-awareness and interpersonal communication." In: B. Dervin; M. Voigt, Eds. *Progress in Communication Sciences Vol. II.* Norwood, NJ: Ablex; 1980:2-49.
48. Bernard, H.; Killworth, P.; Sailer, L. "Informant accuracy in social network data IV." *Social Networks.* 2:191-218; 1980.
49. Nisbett, R.; Wilson, T. "Telling more than we can know: Verbal reports on mental processes." *Psychological Review.* 84:231-259; 1977.
50. Shweder, R. (Ed.) *New Directions for Methodology of Behavioral Science: Fallible Judgement in Behavioral Research.* San Francisco: Jossey-Bass; 1980.
51. Bernard, H.; Killworth, P.; Sailer, L. "Informant accuracy in social network data V." Presented at Sunbelt Social Networks Conference, Tampa, Florida, 1981.
52. Rogers, E. M.; Kincaid, L. *Communication Network Analysis.* New York: Free Press; 1981.
53. Katz, D.; Kahn, R. *The Social Psychology of Organizations.* New York: Wiley; 1966.
54. Reilly, C; Roberts, K. "Task group structure, communication and effectiveness in three organizations." *Journal of Applied Psychology.* 62:674-681; 1977.
55. Richards, W., Jr.; Rice, R. E. "NEGOPY network analysis program." *Social Networks.* 3(3):215-333; 1981.
56. Rogers, E. M.; Agarwala-Rogers, R. *Communication and Organizations.* New York: Free Press; 1971.
57. Tichy, N. "Networks in organizations." In: W. Starbuck; P. Nystron, Eds. *Handbook of Organization Design.* Oxford: Oxford University Press; 1980.
58. Weick, K. *The Social Psychology of Organizing.* Reading, MA: Addison-Wesley; 1969.
59. Rice, R. E. "Resources for longitudinal network analysis." *Connections.* 4(1):10-21; 1981b.
60. Hiltz, S. R. "Controlled experiments with computer conferencing: Results of a pilot study." *Bulletin of the American Society for Information Science.* June, (4):10-11; 1978.
61. Hiltz, S. R. "Using computer conferencing to conduct opinion research." *Public Opinion Quarterly.* Winter: 562-561; 1979a.
62. Hiltz, S. R. "Electronic information exchange-findings." *Telecommunications Policy.* June: 156-161; 1979b.
63. Hiltz, S. R. *The Impact of a Computerized Conferencing System on Scientific Research Communities.* Final Report to the Information Science and Technology Division, National Science Foundation (NSF-MCS-77-27813). New Jersey: Upsala College; 1981.
64. Hiltz, S. R.; Johnson, K.; Rabke, A. "The process of communication in face to face vs. computerized conferences: A controlled experiment using Bales interaction process analysis." *Proceedings of the Association for Computational Linquistics:* 1980.
65. Hiltz, S. R.; Kerr, E. (Eds.) *Studies of Computer-Mediated Communications Systems: A Synthesis of the Findings.* Final Report to the Information Science and Technology Division, National Science Foundation (IST-8018077). New Jersey: Upsala College; 1981.
66. Hiltz, S. R.; Turoff, M. "Evolution of user behavior in computer conferencing systems." *Communications of the ACM.* 24(11):739-751; 1981.
67. Cook, K.; Emerson, R. "Design of a computerized laboratory for the experimental analysis of social interaction." Technical Report, Department of Sociology, University of Washington; 1977.
68. Weiss, B. *Digital Computers in the Behavioral Laboratory.* New York: Appleton-Century-Crofts; 1973.
69. Westin, A. "Home information systems: The privacy debate." *Datamation.* 28(7):100-114; 1982.
70. Westin, A.; Baker, M. *Databanks in a Free Society: Computers, Record-Keeping and Privacy.* New York: Quadrangle Books; 1972.
71. Danowski, J. "An evaluation methodology for computer conferencing: An illustration with a CBBS conference." In: S. Hiltz; E. Kerr, Eds. *Studies of Computer-Mediated Communications Systems: A Synthesis of the Findings*, 1981; 448-480. Also in: M. Burgoon, Ed. *Communication Yearbook 6.* Beverly Hills, CA: Sage, 1982a.
72. Spelt, P. "Evaluation of a continuing computer conference on simulation." *Behavior Research Methods and Instrumentation.* 9(2):87-91; 1977.
73. Penniman, W. *Modelling and Evaluation of On-Line User Behavior.* Final Report to the National Library of Medicine. Dublin, OH: OCLC, Inc.; 1981.
74. Rice, R. E. "Human Communication Networking in a Teleconferencing Environment." Unpublished Ph.D. dissertation, Institute for Communication Research, Stanford University, 1981a.
75. Wynn, E. "Office Conversation as an Information Medium." Unpublished Ph.D. dissertation, Department of Anthropology, University of California, Berkeley, 1979.
76. Penniman, W. "Suggestions for systematic evaluation of on-line monitoring issues." *Proceedings of the American Society for Information Science.* 14:65; 1977.
77. Penniman, W.; Perry, J. "Tempo of on-line user interaction: Compendium of papers." *Proceedings of the Mid-Year Meeting of the American Society for Information Science.* 10-23; 1976.
78. Mittman, B.; Dominick, W. "Developing monitoring techniques for an on-line information system." *Information Storage and Retrieval.* 9(6):297-307; 1973.
79. Kennedy, T. "Some behavioral factors affecting the training of naive users of interactive computer systems." *International Journal of Man-Machine Studies.* 7:817-834; 1975.
80. Penniman, W. "Rhythms of Dialogue in Human-Computer Conversation." Unpublished Ph.D. dissertation, Ohio State University, Columbus, OH, 1975a.
81. Penniman, W. "A stochastic process analysis of on-line user behavior." *Proceedings of the American Society for Information Science.* 12:147-148; 1975b.
82. Chapman, J. L. "A State Transition Analysis of Online Informa-

tion Seeking Behavior." *Journal of the American Society for Information Science.* 32(5):325-333; 1981.
83. Rice, R. E.; Paisley, W. "The green thumb videotext experiment: Evaluation and policy implications." *Telecommunications Policy.* 6(3):233-236; 1982.
84. Warner, P.; Clearfield, F. *An Evaluation of a Computer-Based Videotex Information System for Farmers: The Green Thumb Project.* Lexington: University of Kentucky, Department of Sociology; 1981.
85. Atherton, P.; Cook, K.; Katzer, J. "Free text retrieval evaluation." Syracuse University School of Library Science Report TADC-TR-72-159, Syracuse, NY, July 1972.
86. Lucas, H. "The use of an interactive information storage and retrieval system in medical research." *Communications of the ACM.* 21(3):197-205; 1978.
87. Conrath, D. W. "Communications environment and its relationship to organizational structure." *Management Science.* 20(4):586-603; 1973.
88. Edwards, G. "Organizational impacts of office automation." *Telecommunications Policy.* 2(2):128-136; 1978.
89. Irving, R. "Computer-assisted communication in a directorate of the Canadian federal government." In: M. Elton; W. Lucas; D. Conrath, Eds. *Evaluating New Telecommunications Systems.* New York: Plenum; 1978: 455-469.
90 Johansen, R.; McNeal, B.; Nyhan, M. *Telecommunications and Developmentally Disabled People: Evaluations of Audio Conferencing, Personal Computers, Computer Conferencing, Electronic Mail.* Menlo Park: Institute for the Future, Report R-50; 1981.
91. Leduc, N. "Communicating through computers: Impact on a small business group." *Telecommunications Policy.* 3(3):235-244; 1979.
92. Lippitt, M.; Miller, J.; Halamaj, J. "Patterns of use and correlates of adoption of an electronic mail system." In: J. Hershauer, Ed. *12th Annual Proceedings of the American Institute for Decision Sciences,* Las Vegas, November 1980; 1:195-197.
93. Messier, C. "A computerized electronic mail system." *The Office.* 88(5):82-87; 1978.
94. Meyer, N. "Integration in office automation." *Systems, Objectives, Solutions.* 1(4):217-223; 1981.
95. Miller, J.; Nichols, M. "An update in a longitudinal study of the adoption of an electronic mail system: The effects of system use." In: G. Reeves; J. Sweigart, Eds. *13th Annual Proceedings of the American Institute for Decision Sciences,* Boston, November 1981; 1:218-220.
96. Plain, H. "The impact of computer-linked communication on the transmission of operational and regulatory information in organizations." West Lafayette: Purdue University. Presented to International Communication Association, Philadelphia, 1979.
97. Plain, H. "Computer-linked memo systems: Expectations and utilizations—A case study." West Lafayette: Purdue University. Presented to International Communication Association, Acapulco, 1980.
98. Rice, R. E.; Case, D. "Electronic mail in the university." *American Society for Information Science Proceedings.* 228-230; 1981.
99. Sorg, S. "The acceptance of a computer-based office system in the management area." *AFCET-SICOB "Bureautique" Congress Proceedings,* Paris, 1981; 467-478.
100. Tapscott, D. "Investigating the electronic office." *Datamation.* 28(3):130-141; 1982.
101. Vanlommel, E.; de Brabander, B. "The organization of electronic data processing activities on computer use." *Journal of Business.* 48(3):391-410; 1975.
102. Welsch, L. "Using electronic mail as a teaching tool." *Communications of the ACM.* 23(2):105-108; 1982.
103. Danowski, J. "Perceived impact of computer-mediated communication on other organization communication variables." Presented to International Communication Association, Boston, 1982b.
104. Johansen, R.; Vallee, J.; Spangler, K. *Electronic Meetings: Technological Alternatives and Social Choices.* Reading, MA: Addison-Wesley; 1979.
105. Infomedia. *An Introduction to Infomedia.* Palo Alto, CA: Infomedia Corporation; 1979.
106. Pacanowsky, M.; Hardy, A. "An approach to overtime analysis of communication networks: A computer conferencing example." Presented to International Communication Association, Chicago, 1978.
107. Rice, R. E. "Longitudinal system structure and role occupancy of research groups using a computer conferencing system." In: M. Burgoon, Ed. *Communication Yearbook VI.* Beverly Hills, CA: Sage; 1982.
108. Rice, R. E.; Johnson, B.; Rogers, E. "The introduction of new office technology." *AFIPS Office Automation Conference Proceedings,* San Francisco, 1982; 653-660.
109. Bales, R. *Interaction Process Analysis: A Method for the Study of Small Groups.* Reading, MA: Addison-Wesley; 1950.
110. Cappella, J.; Streibel, M. "Computer analysis of talk-silence sequences: The FIASSCO system." *Behavior Research Methods and Instrumentation.* 11:384-392; 1979.
111. Blau, P. *Inequality and Heterogeneity: A Primitive Theory of Social Structure.* New York: Free Press; 1977.