INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again – beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
HARRISON, William Thomas, 1942-
THE DEVELOPMENT AND APPLICATION OF AN AUDIO TAPE
EVALUATION METHODOLOGY FOR THE ARIZONA
DIAL-A-TAPE MEDICAL INFORMATION SYSTEM.

The Ohio State University, Ph.D., 1974
Education, general

Xerox University Microfilms, Ann Arbor, Michigan 48106

© 1975

WILLIAM THOMAS HARRISON

ALL RIGHTS RESERVED
THE DEVELOPMENT AND APPLICATION OF AN

AUDIO TAPE EVALUATION METHODOLOGY FOR THE

ARIZONA DIAL-A-TAPE MEDICAL INFORMATION SYSTEM

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By

William T. Harrison, B.A., M.A.

* * * * *

The Ohio State University
1974

Reading Committee:

John C. Belland
Kathryn T. Schoen
I. Keith Tyler

Approved by

Adviser

Department of Curriculum and Foundations
College of Education
ACKNOWLEDGMENTS

Almost any research study involves the assistance and cooperation of a number of people in addition to the named investigators. A dissertation effort such as this one, carried out in addition to full-time employment and at a distance of two thousand miles from the student's "home" institution, places considerable added demand on his family and his colleagues. A large group of individuals contributed in many ways to this research project; a sincere debt of gratitude for their valuable assistance is acknowledged to:

Mr. James Barry and Ms. Eleanor Johnson of the Arizona Medical Center Library who provided much assistance and advice in gaining access to the records, the tapes and the people involved in the Arizona Dial-A-Tape project;

Messrs Jim Dick, Jim Peebles and Rod Lorey of the Arizona Medical Center Data Processing Division for access to and considerable assistance in the use of computer facilities;

Mr. John Gaines, also of the Data Processing Division, a most knowledgable, patient and thorough counselor who was extraordinarily helpful in the design, implementation and interpretation of the entire project;

Mr. Jim Coons, a colleague and friend who provided a good deal of assistance and support in several phases of the data compilation;

Dr. Bernard Revisin, University of Arizona College of Medicine Office of Medical Education, who gave much advice and encouragement in the design and implementation of this study;

My dissertation committee, including Dr. Kathryn T. Schoen and Dr. I. Keith Tyler who provided counseling, assistance, and critique along with my adviser Dr. John C. Belland who also provided continuing support and encouragement by letter and telephone during the low as well as the high points of a long, distant effort; and especially

My wife, Gail, who provided helpful encouragement and critical review and who cheerfully endured the numerous family disruptions and sacrifices frequently required over many months to complete this project.
Also unnamed but appreciatively recognized here are the many nurses and physicians from all over Arizona who gave significant amounts of their time in the project, along with the members of my staff at Medical Audiovisual Services, University of Arizona Medical Center, who helped in many small and not-so-small ways, and who shared and endured my preoccupation, impatience, exhilaration, and/or frustration about this study for over two years.
VITA

October 27, 1942 .......... Born — Pomona, California

1965 ..................... B.A., Michigan State University, East Lansing, Michigan

1965—66 ................. M.A., Michigan State University, East Lansing, Michigan

1966–70 ................. Assistant Professor, Television-Radio, Ithaca College, Ithaca, New York

1971–72 ................. Administrative Associate, Instructional Media Section, College of Optometry, The Ohio State University, Columbus, Ohio

1972–present .......... Assistant Director, Medical Audiovisual Services, Bureau of Audiovisual Services, University of Arizona, Tucson, Arizona

FIELDS OF STUDY

Major Fields: (MA) Educational Broadcasting
Broadcast Education
(PhD) Educational Communications in the Health Sciences
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION, THE PROBLEM, &amp; HYPOTHESES</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Continuing Education Needs in the Health Sciences</td>
<td>1</td>
</tr>
<tr>
<td>The Use of Audiotape and the Telephone for</td>
<td>2</td>
</tr>
<tr>
<td>Continuing Education</td>
<td>3</td>
</tr>
<tr>
<td>Dial-Access Audiotape Libraries in Continuing Education</td>
<td>6</td>
</tr>
<tr>
<td>The Evaluation of Dial-Access Information Systems</td>
<td>10</td>
</tr>
<tr>
<td>Review of Related Communications Literature</td>
<td>14</td>
</tr>
<tr>
<td>Rationale for the Study</td>
<td>15</td>
</tr>
<tr>
<td>Problem Statement</td>
<td>15</td>
</tr>
<tr>
<td>Objectives and Hypotheses</td>
<td>16</td>
</tr>
<tr>
<td>Research Objectives</td>
<td>18</td>
</tr>
<tr>
<td>Research Hypotheses</td>
<td>19</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>19</td>
</tr>
<tr>
<td>Assumptions and Delimitations</td>
<td>19</td>
</tr>
<tr>
<td>Summary</td>
<td>20</td>
</tr>
</tbody>
</table>
## II. DESIGN AND METHODOLOGY

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>22</td>
</tr>
<tr>
<td>Description of Dial-A-Tape System</td>
<td>22</td>
</tr>
<tr>
<td>Design of the Study</td>
<td>24</td>
</tr>
<tr>
<td>Methodology of the Study</td>
<td>30</td>
</tr>
<tr>
<td>Call-back Measurement of the Tapes</td>
<td>30</td>
</tr>
<tr>
<td>Time Period of the Study</td>
<td>30</td>
</tr>
<tr>
<td>Processing the Log Data</td>
<td>31</td>
</tr>
<tr>
<td>Analysis of the Log Data</td>
<td>33</td>
</tr>
<tr>
<td>Tape Evaluation Questionnaire</td>
<td>34</td>
</tr>
<tr>
<td>Rationale for the Procedure</td>
<td>35</td>
</tr>
<tr>
<td>Development of the Questionnaire</td>
<td>36</td>
</tr>
<tr>
<td>Selection of the Sample</td>
<td>40</td>
</tr>
<tr>
<td>Selection of the Tapes to be Evaluated</td>
<td>42</td>
</tr>
<tr>
<td>Administration of the Questionnaire</td>
<td>43</td>
</tr>
<tr>
<td>Analysis of Results</td>
<td>44</td>
</tr>
</tbody>
</table>

## III. RESULTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call-back Results</td>
<td>46</td>
</tr>
<tr>
<td>Selection of Tapes for Further Evaluation</td>
<td>51</td>
</tr>
<tr>
<td>Questionnaire Results</td>
<td>52</td>
</tr>
<tr>
<td>Questionnaire Interaction Analysis</td>
<td>60</td>
</tr>
</tbody>
</table>

## IV. CONCLUSIONS, DISCUSSION AND RECOMMENDATIONS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>63</td>
</tr>
<tr>
<td>Conclusions</td>
<td>66</td>
</tr>
<tr>
<td>Discussion</td>
<td>67</td>
</tr>
<tr>
<td>Recommendations for Further Research</td>
<td>70</td>
</tr>
<tr>
<td>Recommendations for Design and Implementation of a Dial-Access System</td>
<td>71</td>
</tr>
<tr>
<td>Summary</td>
<td>72</td>
</tr>
</tbody>
</table>
APPENDIX

A. DIAL-A-TAPE CALL LOG SHEET ......................................................... 74
B. PRELIMINARY DRAFT OF THE QUESTIONNAIRE .............................. 76
C. QUESTIONNAIRE AS ADMINISTERED TO SAMPLE OF HEALTH PROFESSIONALS ............................................................ 80
D. "ADVANCE" LETTER SENT TO SAMPLE ........................................... 84
E. COVER LETTER ACCOMPANYING QUESTIONNAIRE PACKETS ........ 86
F. FOLLOW-UP LETTER SENT TO SAMPLE .......................................... 88

BIBLIOGRAPHY ...................................................................................... 90
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Number of Callers and Rate of Call-back</td>
<td>47</td>
</tr>
<tr>
<td>2. Incidence of Call-back by Tape Number</td>
<td>48</td>
</tr>
<tr>
<td>3. Tapes Selected for Evaluation by Professional Panel</td>
<td>52</td>
</tr>
<tr>
<td>4. Sample Size and Rate of Questionnaire Return</td>
<td>53</td>
</tr>
<tr>
<td>5. Questionnaire Factor Analysis</td>
<td>54</td>
</tr>
<tr>
<td>6. Tape Evaluation: Factor Analysis of Questionnaire Items</td>
<td>55</td>
</tr>
<tr>
<td>7. Questionnaire Items Loading Into Each Factor</td>
<td>56</td>
</tr>
<tr>
<td>8. Pearson Correlation of Questionnaire Factors</td>
<td>58</td>
</tr>
<tr>
<td>9. Tape Scores: Call-back and Questionnaire Results</td>
<td>59</td>
</tr>
<tr>
<td>10. Pearson Correlation: Call-back Scores with Questionnaire Scores</td>
<td>60</td>
</tr>
<tr>
<td>11. Interaction Analysis: Number of Questionnaires Returned by</td>
<td></td>
</tr>
<tr>
<td>Order of Hearing</td>
<td>61</td>
</tr>
<tr>
<td>12. Interaction Analysis: ANOVA of Questionnaire</td>
<td></td>
</tr>
<tr>
<td>Administration Variables</td>
<td>62</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION, THE PROBLEM, & HYPOTHESES

Introduction

A number of diverse efforts are underway to increase the availability to health professionals of up-to-date medical information. One such information system uses short audiotape recordings made by physicians and nurses on a variety of health care topics. These recordings are made available to qualified health professionals via the normal telephone system, usually on a state-wide basis. Such a system exists in Arizona and is known as "Dial-A-Tape". This research investigation was concerned with the development of a means of evaluating individual tapes within the Arizona "Dial-A-Tape" library.

Continuing Education Needs in the Health Sciences

The nearly overwhelming need for practitioners in the health sciences to keep up with advances in their field is well documented [Miller & Harless, 1970]. The need for information and learning programs to maintain professional competencies and state-of-the-art awareness is particularly acute for the practicing physician and his allied health colleagues for several reasons: the competing demands for their time are great; the quantity of information which is new or changing is tremendous; and the urgent need to acquire this information is obvious as they confront daily health care problems of great significance to individuals and communities.

Pressure to use his time most efficiently and effectively is a daily fact of life for the physician. This is significant in the design of educational materials. Typical of this point of view is a Regional Medical Program staff member's observation that:
It is imperative that educators make the continuous educational process... easy, desirable, and immediate; those who are actually involved in health care tend to become overcommitted and despair for time [Hickey, 1971, p.78].

Effective continuing education services must also be very up-to-date; while traditional printed matter is of major importance, the "lag" time from the moment of discovery until the publication and distribution of new information can be quite long.

The overwhelming quantity of information available makes it imperative that continuing education resources allow for selective participation; the health professional simply cannot expect to "return to school" for a broad spectrum update. He must be able to know about and select those learning and informational resources which specifically apply to his current work. Thus, reference systems with sophisticated indexing and compartmentalization of information, made available for his retrieval on demand, are essential.

And finally, easy accessability of continuing education resources must be provided. The mass-distributed and readily-consulted professional journal accounts for the most widely used and rapidly digested source of such learning; it is accessible in almost any setting at any time. Any learning resource which can be accessed at the convenience of the user no matter where he might be, will be of great interest in health professionals' continuing education services:

A real problem involves getting the available information quickly to those who need it most... The problem is or can be particularly serious in many rural and semi-rural areas and in community hospitals that may be relatively isolated from the mainstream of medical advances [Wood, 1972, p.603].

The Use of Audiotape and the Telephone for Continuing Education

Many techniques have been employed to meet the continuing education needs of health professionals. These include individual and group instructional methods, conferences and symposia, and various applications of communications media technology. Intense interest exists in the use of communications media because of their ability to satisfy some of the above criteria for effective continuing education. Included
in media technology applications are the use of films, television, various audio systems, and computers.

Communication by voice is one of our principal means of transmitting information. While extensive use is made of visual material in the health professions, the exchange of purely verbal information is a dominant mode of instruction. A variety of audio systems, therefore, have been applied to medical continuing education, including audiotapes, such as the “Audio Digest” service [Oakley, 1966], and radio broadcasts such as the Ohio Medical Education Network services from Ohio State University [Pace, et al., 1972]. Also used in a variety of ways is the telephone system, the most extensive and readily available system of all for voice communication.

Formal efforts at telephone consultation services include “Mediphone”, a fee-for-service consultation system by phone for emergency use on a national scale. “Mediphone” was implemented by the American Society of Contemporary Medicine and Surgery in 1972 [Bellows, 1972]. More widely reported has been the “MIST” (Medical Information Service, Telephone) program at the University of Alabama Medical Center [Klapper, et al., 1970, 1972]. This is a structured telephone consultation service linking a physician anywhere in the state who has need to consult on a specific problem with an appropriate specialist at the University’s medical center. A similar system of consultation for a particular specialty subgroup, Pediatricians, was instituted in Southern California by the American Academy of Pediatrics [Covell, et al., 1973].

Various other programs have involved the use of the telephone to link groups in formal conference with other groups at different locations or with an expert presenter at a distant point. This is commonly known as the “Tele-Lecture” arrangement [Punwar, 1972].

Dial Access Audiotape Libraries in Continuing Education

One major kind of telephone information service is a marriage of the telephone with an audiotape library resource. Such a system, known by various labels such as “Dial-Access” or “Dial-A-Tape”, was first developed for the use of health professionals by
the University of Wisconsin’s postgraduate medical education staff in 1968 [Meyer, et al., 1970]. Since then, a number of other groups have provided a similar service, in most cases using tapes originally produced in the Wisconsin program.

Basically a dial-access system is composed of a library of short, single-concept audio recordings on a variety of health care topics. These recordings are played on demand over the telephone for health professionals in a given geographic area. (This is not to be confused with those telephone consultation services which provide access to a “live” professional consultant, such as various drug information or “Dial-A-Dietitian” services.)

Tapes have usually been produced through the voluntary efforts of physicians and nurses associated with a teaching medical center; various kinds of professional assistance are normally provided to the author/producer by the staffs of Regional Medical Program units and/or continuing education offices [Meyer, et al., 1970; Berglund, et al., 1969; Wood, 1971; Hickey, 1971; Niles, 1969]. This method of information service has been adopted by a number of medical groups, and it has been adapted to a variety of health-related educational purposes.

Wood has reported in one of the more thorough summaries of information services available to the practicing physician that:

By May, 1970, Dial-Access systems were operating in at least 18 states, telephone consultation services in nine states and telephone conference facilities in six. These figures are conservative, and in all probability will be revised upward on further investigation [Wood, 1971, p.147].

Several health professions groups, other than physicians, are now being served by dial-access systems. The Wisconsin system was expanded to include a separate library of tapes for nurses [Niles, 1970], and many of those tapes are now incorporated in tape libraries around the country, along with copies of the original Wisconsin Regional Medical Program physician-oriented tapes.

The dial-access information concept has also been adapted to serve in-patients with specific health-care information at various hospitals [Bartlett, et al., 1973]. It has also been increasingly applied to general public education programs on topics related to health care. These include the “Dr. Dial” program in Casper, Wyoming [Weiss, 1972] and the
more extensive and widely reported "Tel-Med" project for laymen in the San
Bernardino—Riverside communities in Southern California [Harper, 1972]. Both of these
efforts are reported as highly successful, at least in terms of the number of calls to the
systems.

A recurring theme in the discussions of dial-access services designated specifically
for the practicing health professional is the belief that it is of special significance for the
rural or otherwise isolated person:

There are areas of the country with low population density and few
physicians where, because of the demands on his time and for other
reasons, the practitioner is unable to avail himself of the conventional
modalities of continuing education. Dial-access systems can serve a
useful, if limited, function in providing rapid, up-to-date diagnostic
and therapeutic information to such physicians [Wood, 1971, p.151].

The breakdown of the medical calls we have received shows the service
to be one with appeal to general practitioners who are comparatively
isolated from ready utilization and who have less time to utilize more

Regarding the Nursing Dial-Access program in Wisconsin, Niles [1969] has stated that "it
is primarily intended to meet the needs of the nurse who is practicing in an isolated
setting — the public health nurse, the nurse in a rural hospital or nursing home. . ." [p.329]. Or put another way, Alabama's "MIST" system is valued highly because "by
removing the barrier of geographic distance between colleagues, it brings physicians over
the whole state into the mainstream of medicine" [Klapper, et al., 1971, p.447].

The Arizona Regional Medical Program, faced with a particularly significant
gerographical problem of rural isolation for many health professionals, also established a
dial-access information service. This system, known in Arizona as "Dial-A-Tape", has
been in existence since July, 1970, as a quasi-dial-access system patterned directly after
the pioneering efforts of the University of Wisconsin. A library of tapes purchased from
the Wisconsin Regional Medical Program form the informational content component of
"Dial-A-Tape" in Arizona. The system is housed at the University of Arizona Medical
Center in Tucson, and an auxiliary terminus with its own tapes and playback equipment
is located in Phoenix. It is designated to service physicians and nurses throughout the
state. It is data from this particular operation which formed the basis for this
investigation.
The Evaluation of Dial-Access Information Systems

In order to determine the success of any educational effort, and particularly in order to effectively improve it, a careful scheme of evaluation must be implemented. The evaluation of continuing education services in a meaningful way is not an easy task, however. Wood summarizes the situation for the kinds of information systems being discussed:

The effectiveness of information services as a modality of continuing education is extremely difficult to evaluate. . . . Few attempts have yet been made to determine their impact on physicians or patients. Physicians' needs and habits of seeking information are not well known. They reflect population density of the region, size of the physician's own community, the physician's access to a major medical center, the physician's age, continuing professional education organizations. . . . Information services are badly needed in many rural and semi-rural areas. Many information services are not being utilized to full advantage, but the reasons for this are as yet unclear [Wood, 1971, p.151].

The significant item to measure, of course, is the degree to which such continuing education efforts impact upon the quantity and quality of patient care; any other measure taken has as its purpose a more expedient, if less direct, way of estimating that ultimate result. To directly determine such an impact is virtually impossible; the multiplicity of factors affecting patient care is so diverse and uncontrollable for assessment purposes that information provided by system users must be obtained to enable estimates of effectiveness and as a guide to change. Most evaluation of dial-access systems includes measures along a number of dimensions of the project which relate to size and amount of use of the resource. Subjective judgement by users of the services has also been sought through the common method of short questionnaires sent to a sample of users.

In most systems, a call from a user is first intercepted by a "live" staff member, rather than by an automated system. Therefore, a good deal of data can be quickly obtained and recorded in a log. In virtually all cases, the tape requested, the date and time of day, and usually the location of the caller are noted. In many projects, some or all of
the callers are asked their name and/or professional status and specialty. With this kind of
information, profiles on users’ locations, professional classification and even specialty,
etc., and on rates of use during certain time periods, days, and over periods of time, are
obtained. Geographical profiles are of seemingly special concern to many project sponsors
[Berglund, et al., 1969; Meyer, et al., 1970; Cline, 1971; Florin & Harkness, 1970]. This
may reflect a general evaluative emphasis of Regional Medical Program projects; it is also
relatively easy to do.

The approach taken with the Nursing Dial Access program in Wisconsin is quite
typical:

The evaluation procedure used . . . includes tabulation of the number
of calls received, the tapes requested, the time the calls are received,
and the cities from which calls originate. . . . A sample group of
callers are [sic] asked to provide their names and addresses so they
may be sent postal questionnaires requesting additional information
relating to place of employment, occupation, value of service to caller,
suggestions for additional topics, and comments. When sufficient data
have been obtained, a profile of users will be developed so that tapes
can be prepared to more accurately meet the needs of the callers
[Niles, 1969, pp.335-6].

Another example is the evaluation of the New Jersey dial-access system. The data
were obtained with a mailed questionnaire. The elements covered included: interest in
subject matter, hours of service, professional specialty interest and geographic
distribution of calls. Also, a statement of incidence of use of the system and a general
“useful/not useful” opinion was elicited [Florin & Harkness, 1970]. It is significant to
this study that no specific approach to direct evaluation of individual tapes was
mentioned.

The focus of virtually all evaluation efforts seems to be of a general sort regarding
the “system” — the degree to which it is used, at what times, etc., to what degree
experience with it has been useful, and so on. Of frequent interest and about which much
speculation is raised is the apparent reaction to system promotion. Utilization of each
system is clearly and almost dramatically related to promotional efforts which inform the
potential user population of its availability and contents [Ragatz, 1972; Florin &
Berglund, et al., 1969]. This is stressed in these reports as a key feature upon which utilization of the service seems to hinge.

This emphasis on "system", over concern for quality of elements, is sustained in an overview of information systems provided by Davis [1971]. From the perspective of planning a nationwide dial-access system as part of the National Biomedical Communications network plan, she makes the following observations:

The analysis of the data from the Wisconsin Dial Access System has shown that utilization or demand experience can be explained reasonably well in terms of three basic variables: promotion policy (number of promotions or reminders per year), the size of the reachable population, and library size measured in number of messages. Given that the proposed center is to serve the nation as a whole, population must essentially be taken as given. This leaves two major policy or design variables which can be varied to affect utilization, promotion policy and library size [p. 17].

In many evaluation reports, popularity of those tapes most frequently used is cited. In some cases it is pointed out which few tapes meet a high proportion of the demand. It is possible, however, that these tapes represent only topics for which information is most frequently desired; perhaps the titles of some are more appealing or intriguing for the first-time user who is trying out the system. For example, one tape entitled "Marriage on the Rocks" is reported as among the most frequently selected tapes; another is "Drug Abuse by Teenagers" [Meyer, et al., 1969, p.110; Cline, 1971, p.3]. The data analyses do not preclude the possibility that the information provided on these tapes may, in fact, be of little use. The possibility of superficiality and/or personal, moral biases rather than sound professional judgements are quite likely in a brief overview on topics such as these. If this is the case, their disappointing dimensions to the caller first trying the system may discourage him from calling again.

In many cases, evaluators attempted to determine how many times each users call the system [Meyer, et al., 1970; Berglund, et al., 1969]. This is almost always determined, however, by survey respondents themselves; frequently coupled with this is a statement by the respondent as to his intention to use the system again. This is a significant dimension of evaluation; the incidence of return usage to the tape library as an information resource can be a significant clue to its usefulness to the health care
professional. However, except to obtain general observations and to list frequency counts, the data were apparently not subjected to an analysis of so what?. The need for formative evaluation giving direction for improvement, often remains unfulfilled.

Wood responds with a generally frustrated view about these evaluation efforts:

Numbers of telephone calls and requests for information alone are not enough. Many of the existing programs have yet to be adequately evaluated although most of them have built-in evaluation mechanisms, and the programs are undergoing continuous scrutiny. . . . As to effectiveness, we do not yet have a proper means of evaluating information services [Wood, 1971, p.149].

In a later article, Wood mentions that "despite the interest excited by this [the Wisconsin Regional Medical Program] program, its usefulness for physicians has been difficult to evaluate objectively, and changes both in financing and in professional approach are now under way [1972, p.604].

A systematic, significant plan for evaluating individual tapes is one element of any dial-access system review which seems to have received little attention. Berglund, et al. have provided what may be the most thorough statistical tabulation effort reported on dial-access systems (a study of Minnesota’s participation in the Wisconsin system.) They noted as one of their recommendations that:

Consideration should be given to the continuous assessment of individual tapes. This survey made no attempt to determine to what extent specific tapes were considered helpful by the physicians using them. Some individual tapes, rather than the entire library, may be the cause of the limited discontent uncovered in this study [1969, p.39].

As mentioned earlier, the staff of the New Jersey version of the Wisconsin dial-access service gave no apparent attention to this factor in its evaluation.

Where individual tapes have been discussed, a concern for topics is the focus, rather than for the effectiveness of the actual content or the communication process. In fact, the literature on those systems dealing specifically with information designated for physicians reveals virtually no discussion related to the effectiveness of the audiotape mode of communication per se. Only Hickey, for example, reports the involvement of communications specialists in the development of material for the Texas Regional Medical Program’s Project for Cancer Information [1971].
There are two points in their development when individual tapes can be evaluated in some fashion. One of these is at the time the tape is initially being prepared for use. This is the most common one in the systems studied. The second is at the point of assessing impact of the tape, once it has been accessed and utilized by the intended professional audience. Both kinds of individual tape assessment are considered, in the general context of this investigation, to be of major significance in the management and improvement of dial-access medical information services.

Several of the reports in the literature included discussion of professional expert review panels who critique new tapes; almost always, however, the concern is for the usefulness of the topic and the accuracy of content information. Communications effectiveness appears to be expected rather automatically if the author/presenter is indeed a content expert. This seems to be particularly so when the audience is not expected to range beyond his physician peers. (In those systems mentioned earlier where allied health or lay audiences are anticipated, the expressed concern for communication clarity, lucidity, effectiveness, etc. is much more prominent [Niles, 1970; Bartlett, et al., 1973].) It is seemingly a bias of the medical profession — and one substantiated by the common cartoon character of the dull, uninspiring physician lecturer — that if the content is significant, and if it is of inherent interest to the listener, it is up to the latter to stay alert and make sense out of the presentation; the speaker is held responsible only for the accuracy of the content.

Review of Related Communications Literature

Few persons are charmed by the unchanging expression of the face of a wooden Indian or by the unvaried ticking of a clock. They are no more inclined to attend to a speaker who lacks variety and vitality. The immobilized speaker may simply not understand the basic psychological factors involved in focusing attention and maintaining interest. [Eisenson, et al., 1963.]
Communications effectiveness is clearly something which cannot be ignored, even under conditions of the highly motivated, professionally useful learning situation. The communications field has a long history of study in this area, which was explored for its relationship to the dial-access information system under study.

Speech delivery, a general term relating to the manner in which the speech is presented by the particular speaker involved (apart from the content and structure of the presentation) has clearly been shown to have significant effect on speech effectiveness. In most of the research reviewed, attention has been concentrated upon the persuasive speech, and the measure has been that of change in opinion or belief. However, the range of studies are not solely limited to this type of presentation, and the weight of evidence implies broad interpretation across all manner of informational as well as persuasive speaking.

Heinberg [1963] determined that “delivery is far more influential than is content in determining the general effectiveness” of students’ speeches of self-introduction and of persuasion which were made in speech courses [p.105]. While the motivation for receiving effective communication is undoubtedly much different in this setting than in the dial-access medical information system, it is noteworthy that in his studies “intercorrelations among content, delivery and general effectiveness tend to bear out the analysis obtained through computing regressions in that the delivery-general effectiveness r’s are higher than those between content and general effectiveness” [p.106].

Knower [1936] has conducted more comprehensive studies dealing with changes in attitudes. Of significant interest here is the relationship between attitudes toward the speaker and effectiveness of the speaker’s arguments:

It appears that the subjects who were antagonized by the arguments tended to dislike the speaker or that subjects who were antagonized by a speaker reacted negatively to the arguments. And vice versa, a favorable reaction to arguments tended to produce a favorable judgement of the speaker...

... We find a direct positive relationship between the tendency to give a speaker a good rating and the willingness to admit a change of attitude. ... if subjects liked a speaker they were inclined to admit such a change. That this should be true in spite of the fact that subjects were unable to rate the speakers in accordance with the relative effects they actually produced suggests that we are dealing here with a very general reaction involving speaker, material and auditors rather than a series of specific reactions in the subjects [p.120–1].
If one accepts the idea that the relationship of admitting to a change of attitude in a matter of opinion is closely related to the acceptance of new information or a perception of "usefulness" of the content of an informational presentation, Knower's findings would have rather direct relationship to the dial-access medical information presentation.

Haiman [1949] has made the general observation that "it has long been affirmed by rhetoricians that much of the effectiveness of public speaking is dependent upon the character and personality of the speaker" [p.190]. In his studies to determine if differences in personality and speaking skill would affect the amount of attitude change he concluded that there is a "positive correlation . . . between several factors of ethos and success in the art of persuasion" [p.202]. Ethos was measured in his studies in terms of perceived fairmindedness, sincerity, likableness, competence, and prestige. Andersen and Clevenger [1963] in summarizing experimental research on the effect of ethos report that "the finding is almost universal that the ethos of the source is related in some way to the impact of the message" [p.77].

In many of these studies, concern for ethos includes perceptions about the character of the speaker held by the audience prior to the speech itself. This is not usually a factor in the dial-access medical information tapes; the speakers are essentially "unknowns", although their names are listed in the Dial-A-Tape catalog. However, to the degree that perceptions of ethos obtain while hearing the speaker, these studies have clear implications for the construction and evaluation of such informational tapes.

Riley [1928], Goodrich[1933], Gilkinson, Paulson & Sikkink [1954], and Bowen [1966] have all dealt with the relationship between the skill of the communicator and the effectiveness of his communication in similar ways. The importance of a speaker's ability to organize his presentation, to sound interested, credible, competent, etc. in order to be effective is stressed throughout these studies.

Of specific interest in the presentation of recorded audio material are studies that grow out of the early measures of effectiveness of radio speaking. Two common concerns were the rate of presentation and the quality of voice. Also considered is "conversational" quality.

Ewbank [1932], Lumbey [1933], Cantril and Allport [1935] and Nelson [1948] have conducted studies to determine rates of speaking on radio for optimum recall of
different types of material. Further, Cantril and Allport researched the effects of voice and personality characteristics on listeners. A case for relating studies regarding the effect of ethos on the persuasive effectiveness of public speakers to those dealing with the effectiveness of informational messages presented by "unknowns" via Dial-A-Tape may be made in Cantril and Allport's reporting on the effects of the voice:

Voices have a way of provoking curiosity, of arousing a train of imagination that will conjure up a substantial and congruent personality to support and harmonize with the disembodied voice. Often the listener goes no further than deciding half-consciously that he either likes or dislikes the voice, but sometimes he gives more definite judgments, a "character reading" of the speaker, as it were, based upon voice alone [p.109].

To further support the importance of the voice to effectiveness, they report that "many features of many personalities can be estimated correctly from voice". Of perhaps even greater interest is the extent to which people have common stereotyped perceptions of personality based on voice characteristics. In their study they found that "the uniformity of opinion regarding the personality of a radio speaker is in excess of the accuracy of such opinion" [p.123]. In other words, there is a high degree of agreement among listeners on personality traits revealed through the voice, exaggerated to some extent beyond the real situation. This can mean that the selection of a speaker for a recorded informational message who has any voice characteristics negatively associated in a stereotyped way, even though not in fact accurately reflecting his personality, will be less effective for it.

Specific speaker performance skills are delineated in rating instruments developed for public speaking students by Knower [1929] and for radio speaking by Lawton [1930]. Dimensions of interest to a Dial-A-Tape study include factors of animation, intimate contact, tendency to be "recitative" or to sound as if reading a manuscript, volume, pitch, rate and pronunciation [Lawton] and concerns for affectation, sincerity, rate, engrossment in subject, immediate communication with his hearers, effective holding of attention, and energy and enthusiasm [Knower].

Concern for a conversational approach is expressed by Riley [1928] in the broad spectrum of public address; this has particular meaning today when using the personal medium of the telephone:

One does not have to listen long to the dull, monotonous tones of the novice, the sing-song melody of the spell binder, the sanctimonious tones of the old-fashioned minister, the high sounding meaningless phrases of the
patriotic orator, the colorless academic delivery of the classroom lecturer, or the unthinking absent-minded manner of the memorizer, without his soul yearning within him for a breath of fresh oxygenated conversational speaking [p.233].

He prescribes to the speaker that the listener “must be made to feel that he is on the receiving end of the speech, that he is recognized by the speaker as part of the ‘conversation’. Any kind of detached speaking voice will be unacceptable” [p.240].

In summary form, Bowen [1966] states in his “Reassessment of Speech Delivery” that:

Oral delivery should reflect the speaker’s immediate evaluation of the existing relationships among his thought, his personality, and his listeners. Therefore, we can judge a speaker’s delivery effective if it is instantly appropriate to the content of his message, to his intent, to his “ethos” or perceived personality...[p.21].

Rationale For the Study

The evaluation of specific tapes from a communications effectiveness perspective is considered important for several reasons. The more effective each tape is in its communication mission, the more valuable will be the library service as a whole. No matter how high the motivation to obtain information, perceptions of callers that tapes “aren’t very good” will diminish potential use of the service. And as shown in the earlier communications studies reported above, many human, non-content factors can create “not very good” feelings.

Many parameters of an information system are essentially fixed, once established—at least for a major period of time. But the individual tape is a clearly identifiable, malleable unit. A tape can be added, removed, or changed; it is a manageable unit. Therefore, added emphasis on the effectiveness of individual tapes would seem useful to the user, to the manager, and to the contributing authors of even the smallest systems.

It is also useful to consider individual tapes because there are many dimensions which can be manipulated. Unlike so many variables which are difficult if not impossible to change, there are a variety of factors in producing individual tapes which can be decidedly altered: the organization, scope, pace, technical quality, and the choice of presenter which in turn relates to enthusiasm, attitude, credibility, and so on.
Problem Statement

The purpose of this study, otherwise known as the "research problem", is to devise and test a practical, effective method for assessing the perceived usefulness of individual medical dial-access tapes which will focus on communication effectiveness parameters.

Objectives and Hypotheses

In order to design a more effective evaluation scheme to be applied to the Arizona Dial-A-Tape service, and which subsequently may be of value in similar dial-access systems, some specific value positions and assumptions are set forth below. These are essentially assumed criteria, and they form the basis for the research objectives and the specific hypotheses in this study.

Regarding resource libraries in general:

1. The overall value of the establishment and maintenance of a professional library is greater to the degree that patrons repeat their use of the library's services; i.e. it is of more value if most patrons use it more than once than if each uses it only once. This can be expressed in both economic and professional service value terms. (This does not mean that infrequent, or even one-time use is not valuable, however.)

2. The perceived usefulness of individual items in a library affects one's perception of the usefulness of the library as a whole. Individuals will tend to return to a resource library more frequently if their first use of that library resulted in a helpful or rewarding experience than if it resulted in a wasted or frustrating experience.

3. A valid measure of perceived usefulness of individual materials in a library, therefore, is the degree to which users of these specific materials return (of their own free will) to use other materials in that library.

4. A goal for a librarian, therefore, is to minimize the number of factors in his library system which cause frustration or a sense of futility as well as to maximize those which most nearly fulfill the needs and expectations of patrons.
Regarding dial-access medical information systems specifically:

5. A variety of reasons exist for or against a Dial-A-Tape system user calling back after the first use of the system (hereinafter designated "call-back behavior"). These include such things as access to other information sources, geographical location, elapsed time since medical education experiences, type of patient problems, level of perceived need for continuing education, and the caller's perception of the general usefulness of the Dial-A-Tape system based on his experience with the first-called tape.

6. All of these stimulus variables affecting the call-back choice will be reflected in the mean rate of repeat use, or call-back, after the first call to the dial-access library.

7. These influences on the incidence of call-back behavior are randomly distributed in relation to the first-selected tape, except for the perceived usefulness of that tape.

8. Therefore, a valid measure of the perceived usefulness of an individual tape in the dial-access library is the relative rate at which individuals who first call a specific tape call back again to listen to any other tapes. Specifically, any statistically significant deviation in the call-back behavior of those who first selected a particular tape, as compared to the overall rate of call-back for all tapes, may be attributed to the usefulness of that tape as perceived by those callers.

Research Objectives

Based on the above set of statements, the following objectives were set forth to fulfill the purpose of this study:

1. To measure the relationship between specific, first-selected tapes and the incidence of further calls to the tape library, and to use this data as one means of rating those tapes as an evaluation measure;
2. To develop a comprehensive, minimally reactive rating instrument for a subjective evaluation of individual tapes based on a number of communication variables;

3. To obtain from a representative group of health professionals their evaluation, using the rating instrument, of a sample of those tapes evaluated in step 1;

4. To perform a factor analysis on the rating instrument to determine the significant factors operating in this kind of evaluation of the tapes by health professionals, and the relationship among those factors, and

5. To determine the correlation between the call-back behavior rating and the questionnaire evaluation data for those tapes reviewed in step 3, both in terms of overall rating and on each of the several factors found to exist in the questionnaire.

The attainment of these objectives is expected to provide data on which to base management decisions regarding the deletion or revision of tapes deemed to have significant negative impact on the utilization of this information system and regarding the development of guidelines for the production of new tapes to be added to a library. To the degree a correlation exists between professional judgement and call-back behavior for individual tapes, call-back behavior can be judged as a meaningful indicator which can be monitored routinely as a management decision-making tool to assess the effect of modifications to the tapes and to the system. The relative strength of correlation to call-back behavior of the several rating instrument factors should provide significant insight to the relative importance of particular tape quality dimensions to the callers as well. Also, the rating instrument may be used as a formative evaluation tool for new tapes.
Research Hypotheses

To formulate the statistical research base for this investigation, the following hypotheses are presented:

1. There will be found in the records of all calls to the Arizona Dial-A-Tape service significant deviations from the overall mean rate of call-back behavior of first-time callers for specific tapes, with some tapes obtaining higher and others lower rates of call-back than can be attributed to chance.

2. There will be a positive correlation between deviation from the average rate of call-back behavior for first-time callers of each tape and a rating of quality factors and perceived usefulness of that tape by a panel of representative health professionals.

The hypotheses deal with evidence obtained from the archival data on all calls to Dial-A-Tape as to the existence and validity of indicators of callers' responses to their first experience with the system. If there is significant variation in the quality and perceived usefulness of the tapes, and if this variation affects the degree to which people return to use the dial-access library again, the acceptance of the hypotheses should obtain.

The first statement is presented as a hypothesis, though it deals essentially with a data analysis procedure, because the determination of such a pattern is a major portion of the thesis presented. The second statement will tend to validate the information established in the first as a useful assessment factor for evaluating tapes in the dial-access libraries. Conversely, the first will tend to validate the rating instrument as a useful tool for formative evaluation. Furthermore, the correlative analysis among the survey instrument factors and between those factors and the call-back data generates practically significant information related to tape development and system management.

Note that the assumed criteria and the hypotheses do not indicate that a tape which is used little or not at all is of low value and should be eliminated. The hypotheses will indicate, if accepted, a means for identifying, on a routine monitoring basis, tapes used more than a few times which tend to significantly inhibit further use of the service —
the tapes which "turn people off". To identify these tapes in order to change or eliminate them is in keeping with the value positions outlined earlier. Of course, to identify those tapes which appear to "turn people on" will be useful as well in developing tape production guidelines.

Definition of Terms

First-time caller: A person calling the Dial-A-Tape system during the study period who is a member of one of the health profession classifications included in the study and for whom the Dial-A-Tape service is intended and who has not called the system to listen to any other tape before.

Call-back behavior: The placement of another call to hear any tape in the Dial-A-Tape system by a first-time caller. Call-back failure will be the lack of a call-back during the study period.

First-selected or first-called tape: That tape selected for hearing by the first-time caller.

Professional groups: Categories of health professionals to be considered in this study include physicians (MD's and DO's), nurses (RN's and LPN's), and "others" (including medical and nursing students, allied health staff, and related health professional staff.)

Perceived usefulness: An experience which is professionally rewarding to the individual is considered to have perceived usefulness; specifically, a learning/communication experience which he perceives will in some way prove to have been worth attending to in order to improve his professional competence.

Assumptions and Delimitations

Several assumptions are implicit in the design and conduct of this investigation, in addition to those assumed criteria or theoretical assumptions listed earlier. They are as follows:
1. The reliability of the technical system and personnel service provided throughout the study period is assumed to be uniform, with technical malfunctions, operator error, etc. considered random in relation to time period and to individual tape.

2. The actual distinctions which might be drawn among certain types of physicians, of nurses, and among the components of callers grouped together as "other" are herein ignored. The assumption that there is homogeneity among these groups of Dial-A-Tape callers and among those people serving as professional evaluators is made essentially because of the lack of any alternative; there is simply not an adequate quantity of data to delineate further and assess distinctions within groups.

A qualifying statement is also appropriate at this point regarding the generalizability of any results; this investigation has the purpose of determining in as broad a fashion as possible the feasibility of using call-back behavior to assess users' reactions to a dial-access information library. The user population of the Arizona Dial-A-Tape service includes a broad spectrum of health professionals, with highly urban and remotely rural users included. To the extent the general physician and nursing population of Arizona can be considered representative of the national population or of that of some other region, the results of this investigation are offered as being applicable information in those similar settings for a similarly operated service.

Summary

The need for effective continuing education services for the physician and his health professional colleagues has been outlined. The use of audiotaped information, accessed over the telephone, is shown as one means of serving this need, and it is being applied in several states. Evaluation of these dial-access information systems has been limited and superficial. One specific concern is the evaluation of individual tapes in terms of communication effectiveness.

One dimension of measurement of individual tapes is the degree to which first-time callers of specific tapes call back a second time. This investigation is concerned, therefore,
with the determination of call-back rate for all first-selected tapes, and with the existence of significant variance in this call-back measure. Also of concern is the ascertaining of any relationship between that measure and a survey questionnaire for evaluating some of these same tapes.

The next chapter will detail the relevant literature on the measurement of this kind of behavioral phenomenon. It will also describe the methodology used in the investigation. Chapter Three reports the statistical data resulting from the investigation including the call-back data, the data relating to the administration of the questionnaire and analysis of results, and the correlation between the two. The degree to which these results have meaning for the health professional and the biomedical education and communications professional and recommendations for further investigation are outlined in Chapter Four.
CHAPTER II
DESIGN AND METHODOLOGY

Introduction

The design of this investigation was based on a concern for some of the nearly classic shortcomings of previous attempts to evaluate services such as Dial-A-Tape. A description of the Arizona Dial-A-Tape system is provided first, followed by a discussion of investigative design. The methodology employed is then delineated.

There are two sections to the investigation itself which are suggested by the hypotheses presented in the previous chapter. The first section deals with the measurement of call-back using the archival data in the system call logs, and the evidence suggested in the first hypothesis. The second section is concerned with the development of the evaluation instrument and its administration. Finally, the procedure for determining the relationship between these two kinds of measures is presented; this was the test of the second hypothesis.

Description of Dial-A-Tape System

The Dial-A-Tape system in Arizona consists of a library of over 300 short audio cartridges running from approximately 3 to 8 minutes. Two copies of each of these tapes exist; one is housed in the Arizona Medical Center library in Tucson, and the other is at the Maricopa County Medical Society building in Phoenix. At each site are two Cousina brand tape cartridge playback machines. These tape players are each connected to a separate telephone line. The two telephone lines at each station are on "rotary" so that if the listed number is busy, it automatically steps to the second number for an incoming call.
Upon receipt of a telephone call the staff member answering the phone determines from the caller the number of the tape he wants to hear and some basic identification data. The tape is pulled from the shelf and placed on the tape player. Once it is started, and the volume adjusted if necessary, the staff member hangs up. The tape continues to play through completely for the caller, and it automatically shuts off at its conclusion. The shut-off mechanism also disconnects the telephone line with no further intervention by the staff member. (If the caller wants to hear the tape again or to hear another tape, he must dial again.)

The tapes themselves were acquired from the Wisconsin Regional Medical Program in the cartridge configuration. It was originally anticipated that new tapes would be made locally, but none has been produced to date.

Physicians and nurses in Arizona have been periodically made aware of the service through mailings of the tape library catalog and occasional news releases and stories in various print media. The service originated on July 15, 1970, and catalogs were mailed to all physicians in the state at that time. Later on it was decided that all nurses in the state should receive catalogs as well. These were to be distributed by the State Nurses Association to its membership. However, this plan was never fully carried out, although a substantial number of nurses acquired catalogs through bulk distribution to some hospitals.

In Tucson, the Dial-A-Tape phones are located at the Medical School library circulation desk, and they are answered by employees of the library. This is an ancillary task associated with their main duties involving the circulation and monitoring of library materials. In Phoenix, the Dial-A-Tape phones are answered by the physicians’ answering service operators who are employees of the Maricopa County Medical Society. Physicians and nurses in Maricopa County are told to call the Medical Society number for Dial-A-Tape; callers from anywhere else in the state are directed to call the Tucson number.

If the caller is outside the immediate toll-free area for either station, he is to call “collect”. In Tucson, the Dial-A-Tape phones are part of the University’s centrex system which has toll-free access, on an outgoing basis only, to a state-wide telephone network.
Therefore, collect callers to Tucson are asked their phone number as well as the other basic data discussed below. After eliciting the identification data and tape desired, the caller is told to hang up and the staff member calls him back through the state network. This reduces the cost of a call to the minimum amount rather than for a 5 to 8 minute call. Occasionally this causes a problem because the state network lines can get overloaded during normal business hours and the call may not be returned for several minutes or sometimes even an hour or more. (The negative consequences of this procedure have not been carefully assessed, but it is generally perceived as a significant, though fiscally necessary, hinderance to utilization of the service.)

A log is kept of all calls which includes the name of each caller; professional status, such as MD, RN, Medical Student, etc.; the town he or she is calling from; the date and time of day; the tape number requested; and the caller’s phone number (See Appendix A). One item was added for this study — an inquiry to determine if the caller had ever called Dial-A-Tape before. These items are recorded on a log sheet located next to the telephones and tape players; the sheet is filled out by whomever on the staff happens to answer the phone.

Limitations on this logged data include legibility and accuracy problems, particularly in regard to name spellings. If the library circulation desk was particularly busy — as it often is — or if the answering service lines were busy at the Phoenix station, the Dial-A-Tape logs were sometimes given less than full attention as a consequence. By and large, however, they were substantially complete, and some of the missing data could be reconstructed from other calls by the same person.

*Design of the Study*

Virtually all efforts of instructional innovation comment upon the number of listeners or viewers and their opinions about program worth, but almost none go beyond this superficial assessment of impact [Miller & Harless, 1970, p.25].
Miller and Harless seem to have accurately summarized most of the current approaches to dial-access system evaluation. Both the counting and the soliciting of opinions are superficial. Furthermore, when the general effectiveness of the follow-up questionnaire approach is considered, there may be considerable doubt about the accuracy of even this superficial information. For all the extensive effort and the best of intentions by a number of evaluators, some of the final analyses don't seem well supported by the data.

Few physicians responding to questionnaires about dial-access services have indicated that the information gained via the tapes was of little or no use. Most indicated, conversely, that it was helpful and that they intended to use the system again. Yet, system use tapers off significantly and repeatedly after each promotional effort. And the number of people who call only once is far greater than the expressed sentiment of physician respondents would suggest.

For example, about 80 percent of the physicians surveyed in Minnesota recommended continuation of the service; however, only 46 percent of the callers ever made a second call. Only 27 percent made more than two calls [Berglund, et al., 1969].

It is suggested that the kind of data gathered by mailed questionnaires from physicians regarding dial-access services is less than reliable. A review of the nature of questionnaire data in social science research provides some insight to the problem.

Kerlinger [1964] indicates that people responding to questionnaires “tend to give responses that are socially desirable, responses that indicate or imply approval of actions or things that are generally considered to be good” [p.475]. This idea appears frequently in discussions of the validity and reliability of questionnaire data.

From another group, for example, comes this observation on research methods in social relations:

If people feel they are “guinea pigs” being experimented with, or if they feel they are being “tested” and must make a good impression, or if the method of data collection suggests responses or stimulates an interest the subject did not previously feel, the measuring process may distort the experimental results [Sellitz, et al., 1959, p.97].
Orne [1962] provides some interesting observations on a related phenomenon involving experiment subjects for social science research. He speaks of the “demand characteristics” as perceived by the subject, of the research situation. Subjects want to help the researcher; they hope for him that his thesis is correct. And there is an almost religious response to the “good” of scientific endeavor. They therefore behave “in a manner designed to play the role of a ‘good subject’ or, in other words, to validate the experimental hypothesis” [p.778]. The demand characteristics are those cues which tell the subject what the purpose is — the intended or preferred outcome, or which in some other way give a clue to the hypothesis under examination. It would seem that a questionnaire clearly designed to determine a user’s judgement as to whether an information service is useful presents fairly clear demand characteristics.

A significant aspect of questionnaires vis-a-vis dial-access information services for physicians is that respondents are in fact being asked to make judgements of their colleagues. The tapes are produced by physicians, and the tape system is usually managed by members of the medical profession, often as temporary employees of Regional Medical Programs, or otherwise closely affiliated with them as professional colleagues. This professional relationship may significantly inhibit critical, negative reaction on these questionnaires.

In this regard, Goffman [1959] has presented evidence that supports this notion:

Colleagues may be defined as persons who present the same routine to the same kind of audience . . . Colleagues . . . share a community of fate. In having to put on the same kind of performance, they come to know each other’s difficulties and points of view; whatever their tongues, they come to speak the same social language [p.160].

. . . any social establishment may be studied profitably from the point of view of impression management . . . this will include the conception of one team and of audience and assumptions concerning the ethos that is to be maintained by rules of politeness and decorum. . . . Among members of the team we find that familiarity prevails, solidarity is likely to develop, and that secrets that could give the show away are shared and kept [p.238].

This general, theoretical analysis of social relations is supported in its application to the medical profession by the common understanding that physicians will protect each
other's stature. They do this by strongly avoiding any criticism of each other in any public way; society generally seems determined to sustain the physician's god-like image of near infallibility. This phenomenon exists more strongly in communication across the boundaries of the collegial group than among its members, surely. Perhaps, however, the relationship of the dial-access evaluator is unclear as to colleague status; is he "in" or "out"? This factor may make the reliability of their questionnaires even less predictable.

The most helpful resource for the purpose of exploring the research problem presented is by Webb, et al., [1966] in their work on social science research: *Unobtrusive Measures*. Their writing gave strong direction to and support for the formulation of this study:

The dominant mass of social science research is based upon interviews and questionnaires. We lament this overdependence upon a single fallible method. Interviews and questionnaires intrude as a foreign element into the social setting they would describe, they create as well as measure attitudes, they elicit atypical roles and responses, they are limited to those who are accessible and will cooperate, and the responses obtained are produced in part by dimensions of individual differences irrelevant to the topic at hand [1966, p.1].

They describe these kinds of measures as "reactive" measures. A key concern is that expressed by Kerlinger and others above — the awareness of being tested in some fashion and the biases associated with providing socially acceptable views. In fact, their view of the problems associated with reactive measures is strong: reactive measurement error is "the most understated risk to valid interpretation" of survey data [p.13].

Four classes of reactive error are described. These are the guinea pig effect, role selection, measurement as a change agent, and response sets. The first has been discussed. The second is not so likely to be a problem in dial-access evaluation; the typical respondent from the social scientist's typical sample does not perceive himself in the role of "expert" on a daily basis. Therefore, the man on the street is strongly influenced in his response when he is, as he perceives it, asked to play expert for the researcher. The physician, however, is "expert" by definition. In fact, he can often be observed taking the expert role in areas outside of his training or competency simply because it is so much a part of his daily professional life and therefore his assumed personality.
Measurement as a change agent, the third reactive problem according to Webb and his colleagues, is more likely to operate in dial-access surveys. Whether it is a significant problem for the evaluator, however, is open to question; it depends. So far as eliciting opinion on specific tapes is concerned, if a respondent decides to use the system again, or to react to it differently because he has been interviewed, it does not seem likely to contaminate on-going evaluation of tapes (assuming he is not part of a later sample). It may affect utilization data, but in a positive manner; few dial-access system managers would object to a Hawthorne Effect as a by-product of their evaluative efforts!

The fourth reactive problem, that of response sets, deals with the phenomenon of answering a questionnaire in a manner calculated to give the “right” impression, to be a “good guy”, or the related tendency to simply be agreeable. People generally tend to endorse a statement rather than to disagree with its opposite, for example.

Obtrusive measure problems of this sort can be widely expected in the studies done on dial-access systems. But what can be done to avoid them? Webb and group state that “each of the four types of reactive error can be reduced by employing research measures which do not require the cooperation of the respondent and which are ‘blind’ to him” [p.20]. These measures they call non-reactive or unobtrusive measures, and they direct attention to social science research data not obtained by interview or questionnaire.

One key source of data for such measures, they point out, is the archival record. A great number of research possibilities are brought to mind as their discussion continues:

Governments at various levels provide massive amounts of statistical data, ranging from the federal census to the simple entry of a wedding in the town-hall ledger. Such formal records have frequently been used in descriptive studies, but they offer promise for hypothesis-testing research as well [p.57].

The use of this kind of data for hypothesis testing is subject to limitations, of course. It may have errors of another type which limit its usefulness; selective deposit and selective survival of archival data are mentioned [p.54]. And assumptions made about the relationship of some kinds of data to the behavior under study may seem less direct than in the questionnaire situation. However, the advantages in overcoming reactive problems seem significant:
Besides the low cost of acquiring a massive amount of pertinent data, one common advantage of archival material is its nonreactivity. Although there may be substantial error in the material, it is not usual to find masking or sensitivity because the producer of the data knows he is being studied by some social scientist. This gain by itself makes the use of archival material attractive if one wants to compensate for the reactivity which riddles the interview and the questionnaire [p.53].

One kind of non-reactive measure which relates to dial-access evaluation is the use of library circulation data to show a relationship to other phenomenon. Book withdrawals by library patrons are discussed in relation to a radio book-review program and also to the information-seeking habits of patients compared to different categories of doctors (informative/uninformative) [p.81].

Coming closer yet to a concern with the behavior of physicians, Webb, et al., report in their book on a study of diffusion of innovations among physicians which relates to the use of archival data:

Instead of the more standard, and reactive, tactic of interviewing doctors, they [Coleman, Katz, and Menzel, 1957] elected to go to pharmacy records for information about which doctors prescribed what drugs when. Sampling at intervals over a 15-month period, they related the physician's adoption of new drugs to his social network. Such hardnosed data can be a useful check on interviewing data, provided the effect of collecting such records does not alter the behavior of the record-keepers. This is a very unplausible risk with drug prescriptions, but a reasonable one when dealing with less legally controlled records. The danger is not so much in masking information as it is in improving it [p.93].

Since we are concerned with the actual behavior of physicians in relation to the dial-access information system, to be able to use the actual behavior for evaluation purposes would be, by the above criteria, a major goal. As indicated, however, this kind of unobtrusive measure has its weaknesses as well as the reactive ones. Really, there is a problem with using any single technique alone. As Webb and his colleagues put it:

The principle objection [to reactive measures] is that they are used alone. Interviews and questionnaires must be supplemented by methods testing the same social science variables but having different methodological weaknesses. . . . Once a proposition has been confirmed by two or more independent measurement processes, the uncertainty of its interpretation is greatly reduced. . . . The most persuasive evidence comes through a triangulation of measurement
processes. If a proposition can survive the onslaught of a series of imperfect measures, with all their irrelevant error, confidence should be placed in it. Of course, this confidence is increased by minimizing error in each instrument and by a reasonable belief in the different and divergent effects of the sources of error [pp.1–3].

The focus of this study of the Arizona Dial-A-Tape service is therefore a non-reactive approach to the determination of the perceived usefulness of individual tapes on the basis of actual behavior of system users; this is correlated with a more traditional but carefully designed and administered survey instrument as a means for providing meaningful evaluation information.

Methodology of the Study

Call-back Measurement of the Tapes

The first portion of this investigation was the assessment of the rate of call-back for each first-selected tape in the Arizona Dial-A-Tape system. The log of calls to Dial-A-Tape formed the data base for this process.

Time Period of the Study. This portion of the study involved an analysis of all calls made to Dial-A-Tape from the day it first started, July 15, 1970, until December 31, 1972—a two and a half year period. This time period was selected in order to provide an adequate quantity of data for this study. By including the beginning months, accurate determination could be made of first-time callers and also of all the various categories of adopters vis-à-vis “innovations”. In other words, this time period is considered to more than adequately include those who functioned as the innovator and early adopter types as well as the relatively late adopters in the life of this Dial-A-Tape service.

The study period was divided into two parts. The first ran from July 1, 1970, to September 30, 1972. The second was for three months from October 1, 1972, through December 31, 1972. Data on first-time callers were collected during the entire first period, including first calls and all call-backs. The second part was to determine call-back behavior for those first-time callers calling late in the first study period. That is, if a
first-time call was placed on or before September 30, 1972, the following three month period was monitored to determine if that caller ever called back. This second period, though shorter than the ideal, was deemed adequate time to assess call-back behavior for this investigation. (Approximately 80 to 90 percent of all callers who called back after their initial call did so in less than 90 days.)

Processing the Log Data. The specific interest in reviewing the archival data from the Dial-A-Tape service is the assessment of subsequent call-back activity by first-time callers for each tape. Therefore, a major step was to rearrange the log data from a chronological record to a grouping of all calls made by each individual caller. For each caller, the listing of calls must remain in chronological order so that the first-selected tape may be determined.

Obviously, for a study like this an assessment of call-back activity can be made only if the archival data clearly provides some means of identifying and discriminating among individuals. In this case, it was done by recording the name given by the caller — often the last name only. (A system might use an ID number of some sort instead.) The name was recorded on the log as given, which in the case of many physicians was often simply “Dr. Doe” — no first names or initials. Nurses tended to give first names more frequently as did medical and nursing students and allied health personnel. Nevertheless, in many cases only last names were given, even in these categories.

Individuals with similar sounding or the same names could be separated in many cases by correlation with the town from which the call came or by professional status. Someone named “Jones” from Casa Grande is not likely to be the “Jones” from Flagstaff, for example. Similarly, professional status designations helped to discriminate different but same-named individuals to the extent they were accurately given. Jones, the Tucson RN was assumed as different from Jones, The Tucson MD, for example.

Some allied health people and students were found to identify themselves as MD’s for the first call or two, but in later calls to give what is presumably their real status, such as medical student, or an allied health occupation such as “inhalation therapist”. Where a
group of calls assumed to be associated with the same caller gave different answers for status, the less prestigious designation was taken as the accurate one for purposes of this study.

There was wide latitude in the "spelling of choice" of callers' names on the logs. Continuous familiarity with the entire set of data by one key-punch operator helped to catch errors in spelling, especially with uncommon names which were spelled in different but similar sounding ways, and which by correlation with location, status, juxtaposition of time of the calls, etc., could clearly be judged to be the same person. In some cases it was virtually impossible to tell if a caller was the same person as a previous caller or not.

To the extent that a decision was a toss-up as to whether a given call was from the same person or a new one, it was assumed to be the same person. This was thought to be the more conservative approach so as to minimize false indications of first-selected tapes; if instead, an ambiguous listing were designated as a new, first-time caller, this would create an incidence of a first-call to the tape involved which may have been false. The decision rule employed, however, may have introduced some error in false indication of call-backs. This kind of error would have far lower incidence, however; a choice to associate the ambiguous listing with a previous caller would change the call-back rating of that caller's first-selected tape only when the ambiguous listing was the sole additional call for that person. In many cases, the previous caller had made more than one call anyway. To the extent that this approach to such decisions of interpretation was in error, it was assumed to be random error throughout the data with regard to the first-selected tape and subsequent call-back.

Altogether, 7,153 calls were made during the entire study period (including both segments.) The data on these calls were transferred to punch cards and then read and printed on data processing equipment which resulted in a single print-out of all cards. The cards were then rearranged to provide a listing in which all the calls from a given individual were adjacent in the deck and in chronological order; all caller's names were in alphabetical order.

It had been anticipated at the outset that name spellings, location, and professional status could be verified by the use of state-wide directories of physicians and nurses.
However, this was not possible. The state directory of licensed physicians does not include the great number of resident MD’s practicing as house staff in hospitals throughout the state. It also fails to include MD’s who are associated with the numerous military hospitals in Arizona or with the U. S. Public Health Service Hospitals. These latter, government-based and essentially temporary medical personnel, are not required to obtain state licensure and thus do not appear in any single directory. A complete directory of nurses in the state could not be obtained either for much the same reasons. Therefore, professional status designations were taken at face value from the logs, subject to best-guess decisions described earlier where discrepant data for the same person appeared.

Analysis of the Log Data. The data cards, assembled by hand into the proper order as a chronological file by caller, were subjected to analysis by computer to determine call-back behavior of first-time callers associated with each first-selected tape. The program developed a table of data indicating for each first-selected tape number the number of physicians, nurses, and “others” who did and did not subsequently call back after first calling that particular tape. The total number of first-calls, call-backs, and no call-backs were tabulated for each group and for the entire caller population.

While the computer provided a tally of all this data according to professional grouping, it turns out that there are simply not enough data for any single tape to analyze call-back behavior by professional group. Therefore, further analysis of the call-back information was lumped together for the entire calling population.

All the tapes which received ten or more first-selected calls were used for the final tabulation of call-back scores; those with less than ten were arbitrarily considered by the investigator to have an insufficient quantity of data with which to work.

The total number of call-backs and no call-backs for the entire group of tapes with ten or more calls was calculated. From these figures the mean rate of call-back was determined. The call-back ratio for each tape was then compared against this average rate. The test for significance of any deviation from the mean rate must be compatible with single-sample testing involving means and proportions and a dichotomized nominal scale.
The test used meets these criteria; it is the basic \( z \)-test [Blalock, 1972, p.195].

\[
z = \frac{p_s - p_u}{\sqrt{\frac{p_u q_u}{N}}}
\]

Where:

\( p_s \) = proportion of the sample (proportion of call-backs to no call-backs for a given first-selected tape)

\( p_u \) = predicted proportion, or proportion of the population (the mean rate of call-back for all first-selected tapes accruing ten or more first calls)

\( q_u = 1 - p_u \)

\( N \) = number of S’s in the sample (for each tape)

A \( z \)-score was derived for each tape from the call-back data table in this manner.

The selected alpha level was .05 (two-tailed). This statistical test simply relates each tape’s rate of call-back (proportion of call-backs to no call-backs) to the mean rate for the entire group of tapes and assesses the degree to which deviation from that population mean could have occurred by chance. In this manner, a direct statistical test of the first hypothesis relating to the existence of call-back deviations to a significant degree was applied.

**Tape Evaluation Questionnaire**

While the call-back data evaluation procedure was designed to produce useful information about which tapes are being well received and which are being relatively poorly received, it will not tell why. Part of this study, then, was to develop and administer a questionnaire which deals with various controllable parameters of the development and production of individual tapes and to correlate the results of that questionnaire with the call-back data. Then some meaningful conclusions may be drawn about the practical significance of the call-back information.
A questionnaire-type evaluation instrument, consisting of 17 Likert-scale items was developed. Two groups of physicians and nurses, randomly selected from across the state of Arizona, were asked to call the Dial-A-Tape service and to listen to each of four designated tapes, filling in the questionnaire at the conclusion of each. The returned questionnaires were analyzed for significant factors and for a cumulative rating of each tape on each factor and on overall score. These results were then compared to the call-back z-score of these same tapes to determine the relationship between the two types of measures.

Rationale for the Procedure. In order to overcome previously cited shortcomings of typical evaluation efforts, several considerations were reviewed prior to the design of this component of the study. As pointed out earlier, most formative evaluation of medical dial-access tapes, if it exists at all, is performed by members of the agency which is running the information service project. These people, though thoughtful and well informed in many ways, do not constitute an adequate sample of the target population to determine the effectiveness of the service. A systematically selected sample of the dial-access system's target population was deemed necessary in this study.

The other kinds of evaluation performed have dealt with surveys of those who have decided to use the service; they are almost always follow-up check-off cards trying to determine if the caller found the tape useful, or to determine what prompted him to call, what kind of application it had to his patient care decisions, etc. These respondents have in some sense made a commitment to the system by using it which may bias their reactions to a short follow-up questionnaire. One important feature in the design of this study, therefore, was to select a sample of respondents representative of the entire professional population of the state, rather than from system users or from only the close and available group of medical and nursing faculty.

There would likely be a good deal of variability in the "grading" of tapes among members of a panel of judges, particularly with each judge evaluating only one tape. However, if a group of judges each listen to all of a given group of tapes and respond to questionnaire items on each, their relative assessment of the tapes may be obtained; these
results may be accumulated much more accurately and then correlated with the call-back data more meaningfully.

On this premise, then, a set of questionnaires was sent to a group of nurses and physicians who were each asked to listen to a common group of tapes and then respond to the questionnaire. The results of these questionnaires were carefully analyzed for rating data on the selected tapes. They were also analyzed for interaction among variables in the administration procedure which might render the results less interpretable.

*Development of the Questionnaire.* The development of the questionnaire occurred with several limitations and concerns in mind. Unlike many of the other surveys, there was no specific concern here for the choice or applicability of particular topics on the tapes. Rather, the variables which might be called “communication parameters” were of interest. Given that someone cares to know something on the topic, and assuming he has the “entry level” qualifications for the taped learning experience assumed by the author of the tape content — namely, medical background and related experiences — the concern is for how the tape is qualitatively perceived.

Quality considerations were assumed at the outset to include three general groups of variables over which the producer of a tape has control: the content and the manner in which it is organized, the presentation of the content by the speaker, and the technical quality of the tape. In producing questionnaire items, eight specific categories were considered, amplifying on the above: authenticity of the content, organization of the content, perceived usefulness, technical quality, speaker’s performance, speaker’s credibility, appropriateness of title, and overall rating.

A major consideration in the questionnaire design was the condition under which it was to be administered. Respondents were to be asked to volunteer their time for a project to which they may have had no prior knowledge or commitment. And they were to be asked to evaluate not just one but several tapes. This meant, of course, that they would have to fill out the same questionnaire several times, once for each tape. In order to obtain a useful return on the survey, the length of the questionnaire would have to be shorter than one might like. The number of tapes reviewed would have to be limited for the same reason.
The type of item selected for the questionnaire was a Likert scale item design. This was done because it uses very specific statements with which the respondent may simply agree or disagree by varying amounts. It minimizes the variation in interpretation of the item and simplifies the analysis compared with, for example, a semantic differential or a simple multiple choice item which tends to exclude possible choices of interpretation or response. It is also easy to understand and to respond to quickly.

The items used in the development of the questionnaire were drawn from an item pool. A collection of statements about the various dimensions appropriate to the designated categories was developed over a period of time. A group of about sixty Likert-type statements were written. These were reviewed with the investigator's colleagues and with a few of the hospital staff at the Arizona Medical Center. This informal review dealt with the clarity or meaning of the items; in effect this was done as a means to increase the face validity of each item. Those which turned out to appear confusing or to have more than one interpretation were either rewritten or dropped from the pool. The surviving items, numbering 48, were put on machine-readable cards, one statement per card. These were arranged in random sequence and run through a reader-printer, thus generating a quick and variable set of questionnaires. This draft was administered to six health professions personnel working in the Medical Center. Each was asked to call two tapes and fill out a questionnaire on each. The tapes were one each from the high end and one from the low end of the ranked listings of tapes derived from preliminary call-back data.

The items were then reviewed with each respondent as to clarity, and those groups of items which were variations of asking essentially the same thing were reviewed for internal consistency. The length of the questionnaire was also discussed with each respondent.

From this experience, the pool was shortened and revised to a total of 25 items. This became the basis for the printed questionnaire which was subjected to a reliability assessment (See Appendix B.)

A test-retest approach was used to test reliability. The test draft of the questionnaire was administered to a group of about 65 junior- and senior-level nursing
students attending a course in nutrition during the summer of 1973 at the University of Arizona. While desirable, it was not possible to have these students actually call the Dial-A-Tape service and listen to the two tapes selected for this section of the study on two different occasions; this would have swamped the staff at the Medical Center Library. Therefore, the two tapes were played on a standard cassette player to the class as a group during normal class time. The questionnaires were distributed immediately prior to hearing the tapes. Ample time was given to fill out a questionnaire for the first tape before hearing the second.

One week later the process was repeated. The same two tapes were again played for the whole group in the same order. Identical copies of the questionnaires were distributed prior to playing the tapes, and again ample time was allowed for filling out one questionnaire before hearing the second tape.

Two tapes were used to provide some redundancy in the effort to determine reliability; if two rather different tapes yielded approximately the same results with this group, greater confidence could be placed in the results. Also, a reliability test-retest on one tape might reveal weak items which would not show up for the other, and vice versa, due to the nature of the particular tapes involved.

One shortcoming of this approach is that it does not fully replicate the actual situation under which the questionnaire will be used. While the test-retest format for reliability determination is essentially valid for tests of basic attitudes or general perceptions, it is likely to generate distortion of the results where the tape is heard a second time for a second testing, as in this application. The second hearing and re-testing itself is likely to affect the attitude toward the tape and thus confound a reliability measure.

To the extent this is a problem, it may have been accentuated by the fact that the nursing students were not pleased about participating, particularly when they learned that they would have to repeat the process a second time. (The students were not told, until the beginning of the re-test session that they would be hearing and evaluating the same tapes a second time.)
A total of 106 pairs of questionnaires resulted from the 53 students who were administered the instrument both times. For the purposes of determining item reliability, the pairs of responses for each item were compared in a cross-tabulation process using Goodman and Kruskel's gamma coefficient, a conservative measure of ordinal data. Those items which fell below .500 on the gamma coefficient were considered suspect; these were dropped. (A .500 level for gamma is not to be confused with a .500 result on a Personian r correlation; it is a much “stronger” result than that. Gamma is a measure of association, to measure the reduction in predictive error between one set of rankings and another. See Mueller, et al., [1970], pp. 279-294.)

Correlation coefficients among test items were also determined on the same questionnaire results to assist in narrowing down the item pool for the final draft. Those items which were found to have a very close relationship according to the correlation coefficients were considered as a cluster; in some cases where the coefficient was high, the item with the lower reliability gamma result was dropped.

A factor analysis was also performed on these trial responses of the nursing students. The position of each item within each factor was used as a determination for reducing the number of items; where an item had no strength in any of those factors considered significant (Eigenvalue of 1.000 or greater), it was dropped. When an item showed major strength in one of the factors, it was kept. A cross check of this criteria against the correlation coefficients and on the reliability scores was made so that a sufficient number of items for each factor remained in the questionnaire.

By this set of procedures, eight items were dropped. A total of seventeen items was retained for the finished draft of the questionnaire. These items were then randomly ordered on the final version (Appendix C).

The heading of the revised draft of the questionnaire was modified to include the date and time of the call as well as the number and title of the tape being reviewed and the professional status of the evaluator. The date and time information was requested simply as a means of confirming the order in which the tapes were heard, and to correlate the questionnaires with the log of the calls should they get separated and be difficult to identify in relation to other questionnaires from the same evaluator.
The reduced number of items fit snugly onto two pages. A third page was added for open-ended comments which might provide data and attitudes not elicited through the individual questionnaire items.

Selection of the Sample. The selection of the sample, the panel of evaluators to review the tapes, involved several considerations. These included the sub-groups to be involved, the number of returns needed for adequate statistical analysis, the percentage of return one could expect under these survey circumstances, and the capacity of the Dial-A-Tape system to handle an increased volume of calls.

While the call-back statistics had been evaluated on the basis of three sub-groups—physicians; nurses; and “others,” which included medical and nursing students and a variety of allied health personnel—it was determined that the managers of the service intended it primarily for practicing physicians and nurses. (Those “others” who found their way to the service had done so through inadvertent means, and in fact they were excluded from using the service through the Maricopa County facility; the Tucson facility staff was more liberal in its interpretation of proper access to the service.) Therefore, it was decided that the panel of evaluators would include only the two primary groups of callers, physicians and nurses.

It was also determined that two separate groups of tapes would be evaluated by what would be essentially two distinct sub-groups of the sample, thus providing some redundancy to this element of the study.

In order to provide a sufficient amount of ranking data from the questionnaires to test the validity of call-back phenomenon as an evaluation criterion, it was decided that each evaluator would listen to four tapes. This would require a total of about 60 minutes to call each tape, listen to it, fill out the questionnaire, and prepare them all for mailing. Because this was such a large demand on respondents’ time, a relatively low percentage of return was anticipated; somewhere around 50 percent was anticipated from physicians, and about 70 percent was expected from nurses. These figures were simply educated guesses, principally from several people who have worked with Regional Medical
Programs and have observed the questionnaire process applied several times to health professionals.

Assuming that at least 20 sets of responses from within each professional evaluator group are desirable for solid statistical evaluation for each group of tapes, a minimum sample size of 80 physicians (two groups of 40) and 60 nurses (two groups of 30) was determined. To allow for some error in judgement on the rate of return, and for variability among groups within the sample, these figures were actually set at a sample size of 100 physicians and 75 nurses. While it would have been even more secure to use a larger number, the load on the Dial-A-Tape system itself was a major factor, particularly in relation to the Maricopa County installation.

The physician sample was determined with the use of the current edition of the directory of physicians published by the State Board of Medical Examiners and a table of random numbers. Though there were other MD's practicing as residents or military personnel in the state, they are not included in any available directory and are thus not included in the sample.

First, 75 names were selected by picking two names from each column of the alphabetical listings. The particular two were selected by means of the random number table. While this procedure was considered statistically adequate to represent the population of permanent licensed physicians in the state, it was noted that there is a heavy concentration of physicians in the Phoenix and Tucson areas. With such a small sampling from that directory, very few physicians from non-metropolitan areas were actually included in the resulting sample. Because the impact of this dial-access service on rural personnel was considered important, the sample was purposely distorted to increase the proportion of non-metropolitan individuals. A second group of 25 physicians from non-major city locations (any Arizona community except Tucson, Phoenix and surrounding suburbs, and Flagstaff) was selected, bringing the total sample to 100. This group was selected by again going through the directory, using the random table to select two names from each column. If a selected person met the location criteria, he was added to the sample. This process was repeated about three times through the entire directory until 25 names were generated.
There was no obtainable current, published list of registered nurses in Arizona. However, the State Nursing Association agreed to cooperate in providing names and addresses for a sample. The membership in this organization is voluntary, and it could not be determined directly and statistically to what degree its membership reflects the characteristics of the RN population in the state as a whole. The Director of the Association indicated that the membership did, in her estimation, represent a cross section of this population in a general sense. At least it can be stated accurately that rural and metropolitan, newly graduated and nearly retired, and public health and private duty member categories are all represented in the membership.

The Association provided from its addressograph plates 125 mailing labels representing every twentieth name in the membership file. This resulted in a systematic sample which, if not fully representative in the purest sense, provided a broad cross section of RN's in Arizona to participate in the study. From this listing, 75 were selected, again by using a table of random numbers.

Selection of Tapes to be Evaluated. As mentioned earlier, it was determined that two subgroups of the total sample would be derived and each group would listen to a separate set of tapes. This was to provide some redundancy of application of the basic design. It had also been determined that there would be four tapes in each group. This required, then, a selection of eight tapes for evaluation.

The key interest in the study is to relate the relative scoring of tapes by the evaluators to the scoring derived from the call-back data. Therefore, each set of four tapes was determined by selecting a tape from the extreme high end of the z-score listing for call-back activity, one from the low end, and since an even number had been selected for the group, the two remaining were selected from a z-score position of about +1.0 and from about −1.0. Since the tapes were to be judged by a cross-section of physicians and nurses with presumably diverse backgrounds and interests, tape topics of more general interest were selected where there was a choice of tapes with approximately the same z-score.
Administration of the Questionnaire. A personally addressed and signed letter was sent to each person in the sample one to two weeks before he/she received the questionnaires (See Appendix D.) This letter was to solicit interest and to give general information about the study. The intent was to generate interest and a sense of positive anticipation in the subjects prior to being "hit" with the reality of the four questionnaires and subsequent time demands. A number of colleagues agreed this was a good idea; since the letter was sent to all members of the sample, there is no way to determine the effectiveness of this advance communication.

A sufficient number of packets, each containing four questionnaires, and a stamped return envelope, plus the cover letter (Appendix E) were then assembled. Each letter was again personally headed with the addressee’s name, address and a personal salutation, and the letter was personally signed. Each questionnaire had the tape number and title typed directly onto the line designated for this information. As indicated in the cover letter, a number — 1 through 4 — was marked by hand in red in the upper left corner of each questionnaire designating the order in which the tapes were to be heard. A large return envelope was also enclosed with first-class postage affixed and the return address stamped on. On the back, under the flap of this envelope, “thank you” had been written by hand.

All the elements which related to personalizing, such as individually headed and signed letters, first-class postage, etc., were done to increase the sense of personal communication and the importance of the material. It was hoped this would overcome the negative impact of the large time demands of the material.

The questionnaire packages were mailed over a period of about eight days, starting one week after the “advance” letter. The mailing was staggered so as to minimize the impact on the Dial-A-Tape station answering staff. Fourteen days after the last group of questionnaire packages was mailed, a follow-up letter was sent to those who, according to the logs from Tucson and Phoenix, had not yet made any calls to the service.

The returned questionnaires began to arrive almost immediately, and the last ones came in as long as four weeks later. The items were hand scored, and these scores, along
with other data from the heading on the questionnaire were converted to machine-readable cards. The data included: a four-digit ID number for each respondent, a designation as to which group of tapes the respondent had listened to, the tape number, the order position number (1–4) for that respondent and that tape, the date and time of the call, the professional status (RN or MD), a score for each of the 17 questionnaire items (1–5) and an indication as to whether or not the respondent had made any additional comments on that tape.

Analysis of Results. The questionnaire data was subjected to analysis using the computerized Statistical Package for the Social Sciences (SPSS) from the Vogelback Computing Center, Northwestern University. This package was made available through the University of Arizona’s Computer Center and the Division of Data Processing, Arizona Medical Center.

Each questionnaire item was entered initially in the computer as a separate case. The concern in assessing the questionnaire results is for the relative scoring of the tapes by each individual; however, it was determined that entering each questionnaire as a separate case and processing the statistics accordingly would yield the same information and accommodate the design of the statistical analysis package.

Basic statistics were run on each questionnaire item yielding means, standard deviations, and standard scores. For the factor analyses and the subsequent analyses of variance and correlation determinations, these standard scores were used in place of raw scores for each questionnaire item.

A factor analysis was made to yield the major factors within the 17 questionnaire items. The factor scores and a total score were then compared against several of the variables involved in the testing procedure to detect any significant interaction. This was carried out with the analysis of variance procedure. The variables against which the scores were checked included the tape group (each of the two different groups of tapes/respondents), the order of hearing of the tapes by each respondent, the professional status groups (physicians or nurses), and the incidence of free comments by respondents. Also analyzed for significant interaction were the total and the factor scores for each tape.
A check was also made for interaction among the study design variables which could in some way have had an effect on the results. Included was a cross tabulation of each of the following factors: tape group, tape number, order of presentation, professional status, and the incidence of free comments. The Chi-square statistic was used in assessing the significance of interaction among each pairing of these items.

By this last procedure, the effect of — for example — professional status on the relative rate of return of the questionnaires from each tape group could be assessed. Similarly interaction between the order in which tapes were heard and the rate of return was checked. Also, the degree to which any relationship existed between members of a tape group or professional status group and the incidence of free comments was similarly assessed.

From the scoring data mentioned above, a net score for each major factor, as well as a total score, was determined for each tape reviewed. These results were then correlated by means of the Pearson correlation statistic with the call-back z-scores derived for each tape in the first part of this study. This process yielded the amount of correlation between the call-back behavior and the judgements of the panel of evaluators, thus putting to test the second hypothesis concerned with this relationship.
CHAPTER III
RESULTS

The data resulting from the investigation are presented in two parts, reflecting the nature of the study. The first section reports the data regarding the incidence of call-backs to the Dial-A-Tape service following specific first-selected tapes, as indicated by the Dial-A-Tape logs. The second shows the results of the questionnaire and the relationship of those results to the call-back data.

Call-back Results

The portion of the study regarding the incidence of call-backs included all calls made to Dial-A-Tape between July 15, 1970, and December 31, 1972. The period of time between September 1, 1972, and December 31, 1972, was monitored only for call-backs following first-time calls prior to September 1; no additional first-time callers during this latter period were tabulated.

The total number of calls placed to Dial-A-Tape during the entire study period was 7,153. The total number of tapes selected on first calls was 340.

Shown in Table 1 is the basic data on the number of callers in each of the three professional groups initially designated for study and the level of call-back activity for each. With 2,560 callers selecting from 340 tapes on their first call, a large number of tapes subsequently had very few first-calls placed to them. Because the N was so small in these cases, they were not of value in the analysis. Therefore, the 88 tapes with ten or more first-calls were selected from the call-back data results for further analysis. The rates of call-back for each professional group for these 88 tapes are also shown in Table 1. This comparison indicates the difference, which is not significant, in call-back rate between that group of tapes obtaining ten or more first-calls and the entire group of first-selected tapes.
### Table 1

**NUMBER OF CALLERS AND RATE OF CALL-BACK**

<table>
<thead>
<tr>
<th>First-called tapes</th>
<th>Calling activity</th>
<th></th>
<th></th>
<th>Mean rate of call-back</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First-time callers</td>
<td>N of first-time callers who called back</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Physicians</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All tapes</td>
<td>857</td>
<td>374</td>
<td></td>
<td>.436</td>
</tr>
<tr>
<td>Tapes with first-calls≥10</td>
<td>392</td>
<td>177</td>
<td></td>
<td>.451</td>
</tr>
<tr>
<td><strong>Nurses</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All tapes</td>
<td>1336</td>
<td>480</td>
<td></td>
<td>.359</td>
</tr>
<tr>
<td>Tapes with first-calls≥10</td>
<td>737</td>
<td>255</td>
<td></td>
<td>.346</td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All tapes</td>
<td>367</td>
<td>118</td>
<td></td>
<td>.332</td>
</tr>
<tr>
<td>Tapes with first-calls≥10</td>
<td>209</td>
<td>60</td>
<td></td>
<td>.287</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All tapes</td>
<td>2560</td>
<td>972</td>
<td></td>
<td>.380</td>
</tr>
<tr>
<td>Tapes with first-calls≥10</td>
<td>1338</td>
<td>492</td>
<td></td>
<td>.368</td>
</tr>
</tbody>
</table>

The data matrix for those tapes receiving ten or more first-calls is presented in Table 2. Included is a breakdown, by professional category, for first-calls which were followed by a call-back to Dial-A-Tape and those first-calls which did not result in such a call-back. Even after eliminating the majority of the tapes which accrued less than ten total first-calls, it was additionally necessary to abandon any effort to subdivide the data on the basis of professional status. As may be seen in Table 2, the N's in the professional
<table>
<thead>
<tr>
<th>Tape No.</th>
<th>Call-back data</th>
<th>First-calls followed by a call-back</th>
<th>First calls NOT followed by a call-back</th>
<th>Total first calls</th>
<th>Percent call-back</th>
<th>z-score of call-back</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MD</td>
<td>RN</td>
<td>Other</td>
<td>MD</td>
<td>RN</td>
<td>Other</td>
</tr>
<tr>
<td>255</td>
<td>7</td>
<td>8</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>55</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>501</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>285</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>278</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>206</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>525</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>305</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>105</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>531</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>565</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>198</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>406</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>211</td>
<td>4</td>
<td>13</td>
<td>5</td>
<td>6</td>
<td>17</td>
<td>6</td>
</tr>
<tr>
<td>49</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>553</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>74</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>53</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>241</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>292</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>79</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>169</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>129</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>50</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>533</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>220</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>518</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>256</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>166</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>135</td>
<td>5</td>
<td>6</td>
<td>0</td>
<td>3</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>264</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>218</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Tape No.</td>
<td>First-calls followed by a call-back</td>
<td>First calls NOT followed by a call-back</td>
<td>Total first calls</td>
<td>Percent call-back</td>
<td>z-score of call-back</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------</td>
<td>----------------------------------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td>RN</td>
<td>Other</td>
<td>MD</td>
<td>RN</td>
<td>Other</td>
</tr>
<tr>
<td>236</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>527</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>159</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>34</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>69</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>530</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>567</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>326</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>452</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>453</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>514</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>523</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>529</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>445</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>76</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>29</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>269</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>456</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>44</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>138</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>411</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>112</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>182</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>287</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>299</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>173</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>230</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>410</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>
### TABLE 2, Continued

**INCIDENCE OF CALL-BACK BY TAPE NUMBER**

<table>
<thead>
<tr>
<th>Tape No.</th>
<th>First-calls followed by a call-back</th>
<th>First calls NOT followed by a call-back</th>
<th>Total first calls</th>
<th>Percent of call-back</th>
<th>z-score of call-back</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MD</td>
<td>RN</td>
<td>Other</td>
<td>MD</td>
<td>RN</td>
</tr>
<tr>
<td>544</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>439</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>286</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>42</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>65</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>276</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>33</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>568</td>
<td>1</td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>252</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>328</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>440</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>72</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>93</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>409</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>552</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>232</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>95</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>88</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>298</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>208</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>11</td>
</tr>
</tbody>
</table>

* p < .10
** p < .05
***p < .01
categories are still too small to be of much significance. Therefore, it was not practical to continue data analysis by professional group, and all subsequent results are in terms of the total data for each tape.

The percent or proportion of call-back is shown for each tape, along with the call-back z-score for that tape, based on the mean rate of call-back (.368) of the entire group of tapes with ten or more first-calls. The table is presented in descending order of z-score to most clearly indicate the distribution of results on that statistic.

Because trends are of interest in this study, the alpha level of .10 is indicated as well as the more traditional .05 and .01 levels. The levels of significance are based on a two-tailed analysis.

An analysis of the data in Table 2 represents the test of the first hypothesis regarding the existence of significant deviations in call-back rate for specific tapes in the Arizona Dial-A-Tape log data. There are in fact seven tapes which exceed the .05 level of significance. Thus, the first hypothesis is accepted. It stated that:

There will be found in the records of all calls to the Arizona Dial-A-Tape service significant deviations from the overall mean rate of call-back behavior of first-time callers for specific tapes, with some tapes obtaining higher and others lower rates of call-back than can be attributed to chance.

The incidence of call-backs as recorded by the logs of the service does appear to provide, therefore, a discriminating measure of caller behavior toward the Dial-A-Tape system based on reaction to first-selected tapes.

Selection of Tapes for Further Evaluation

Two groups of four tapes each were selected on the basis of the call-back data for further evaluation by the panel of health professionals. The tapes selected, including title and tape number along with the call-back z-score of each, are shown in Table 3. The first and fourth tapes in each tape group are from the extreme ends of the distribution; the second and third of each represent an approximate position in the distribution of +1.0 and −1.0 respectively. The remainder of the study dealt with a comparative evaluation of
TABLE 3

TAPES SELECTED FOR EVALUATION
BY PROFESSIONAL PANEL

<table>
<thead>
<tr>
<th>Call-back</th>
<th>Tape Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Group 1

| 2.636 | 55 | Angina Pectoris, Present Status & Treatment of          |
| 0.881 | 49 | Delerium Tremens, Management of                      |
| -0.997| 13 | Marriage On the Rocks                                |
| -2.128| 88 | Bee Sting, Treatment of                              |

Group 2

| 2.663 | 255 | Alcohol Withdrawal, Acute, Treatment of          |
| 1.031 | 406 | Drug Abuse & The Role of the Professional Medical Person |
| -0.913| 33  | Suicidal Threat, Recognition & Treatment of        |
| -2.740| 208 | Thyroid Hormones, Use of                          |

these tapes and a correlation of the questionnaire results with the above call-back z-scores.

_questionnaire Results_

The first dimension considered in reporting the results of the questionnaires is that regarding the number of questionnaires distributed, which is to say the size of the sample, and the subsequent number of questionnaires returned. Table 4 presents this information which generally indicates that 34 percent of those physicians requested to participate in the evaluation effort responded, and 66 percent of the nurses responded. There is a significant difference in the rate of return from the nurses between those asked to listen to the first group of tapes and those asked to listen to the second group. The practical
TABLE 4
SAMPLE SIZE AND RATE OF QUESTIONNAIRE RETURN

<table>
<thead>
<tr>
<th>Tape group</th>
<th>Questionnaires</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number distributed</td>
<td>Number returned</td>
<td>Percent returned</td>
</tr>
<tr>
<td>Physicians</td>
<td>Group 1</td>
<td>50</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Group 2</td>
<td>50</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Both</td>
<td>100</td>
<td>34</td>
</tr>
<tr>
<td>Nurses</td>
<td>Group 1</td>
<td>38</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Group 2</td>
<td>37</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Both</td>
<td>75</td>
<td>49</td>
</tr>
<tr>
<td>Total</td>
<td>Both</td>
<td>175</td>
<td>83</td>
</tr>
</tbody>
</table>

meaning of this difference (81 percent versus 51 percent) has not been determined. All other interaction parameters which were measured (see below) have indicated no meaningful differences; it is speculated that any effect other than chance on this dimension must relate to the level of interest the tape titles raised among the nurses in each group.

Each questionnaire item was scored with a point value of from 1 to 5, with 5 being the most positive indication of evaluation of the tapes in terms of the item. In other words, "strongly agree" scored as '5' while "strongly disagree" scored '1'. A negative statement, for which a high level of agreement by the respondent would indicate a poor assessment of the tape, was scored in reverse fashion; the highest level of agreement with the item would cause the lowest score. Thus, every item resulted in a high score being a strong favorable assessment of the tape on that particular characteristic, regardless of the positive or negative nature of the Likert-type statement.

In order to facilitate interpretation of the data, the first analysis of the questionnaire item scores was a factor analysis. The procedure used for this analysis
TABLE 5
QUESTIONNAIRE FACTOR ANALYSIS†

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor Loading*</th>
<th>Factor Loading*</th>
<th>Factor Loading*</th>
<th>Communality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1</td>
<td>Factor 2</td>
<td>Factor 3</td>
<td>Communality</td>
</tr>
<tr>
<td>1</td>
<td>.691</td>
<td>.435</td>
<td>.192</td>
<td>.704</td>
</tr>
<tr>
<td>2</td>
<td>.355</td>
<td>.810</td>
<td>.036</td>
<td>.783</td>
</tr>
<tr>
<td>3</td>
<td>.138</td>
<td>.138</td>
<td>.879</td>
<td>.810</td>
</tr>
<tr>
<td>4</td>
<td>.274</td>
<td>.735</td>
<td>-.125</td>
<td>.631</td>
</tr>
<tr>
<td>5</td>
<td>.199</td>
<td>.720</td>
<td>.356</td>
<td>.685</td>
</tr>
<tr>
<td>6</td>
<td>.771</td>
<td>.416</td>
<td>.223</td>
<td>.816</td>
</tr>
<tr>
<td>7</td>
<td>.839</td>
<td>.172</td>
<td>.039</td>
<td>.734</td>
</tr>
<tr>
<td>8</td>
<td>.235</td>
<td>.778</td>
<td>.135</td>
<td>.679</td>
</tr>
<tr>
<td>9</td>
<td>.698</td>
<td>.284</td>
<td>.129</td>
<td>.585</td>
</tr>
<tr>
<td>10</td>
<td>.839</td>
<td>.078</td>
<td>.121</td>
<td>.725</td>
</tr>
<tr>
<td>11</td>
<td>.331</td>
<td>.515</td>
<td>.182</td>
<td>.406</td>
</tr>
<tr>
<td>12</td>
<td>.760</td>
<td>.442</td>
<td>.212</td>
<td>.817</td>
</tr>
<tr>
<td>13</td>
<td>.724</td>
<td>.296</td>
<td>.031</td>
<td>.613</td>
</tr>
<tr>
<td>14</td>
<td>.146</td>
<td>.023</td>
<td>.882</td>
<td>.800</td>
</tr>
<tr>
<td>15</td>
<td>.285</td>
<td>.809</td>
<td>-.011</td>
<td>.735</td>
</tr>
<tr>
<td>16</td>
<td>.775</td>
<td>.385</td>
<td>.118</td>
<td>.763</td>
</tr>
<tr>
<td>17</td>
<td>.715</td>
<td>.297</td>
<td>.044</td>
<td>.601</td>
</tr>
</tbody>
</table>

†Factors with Eigenvalue ≥ 1.000
*Varimax rotated factor matrix after rotation with Kaiser normalization

involved the production of a varimax rotated factor matrix, and the results are shown in Table 5. The rotation procedure is a standard data transformation that clarifies relationships among the components; it is in fact a form of principal components analysis.

As is standard in a principal components analysis, only those factors with an Eigenvalue of 1.00 or greater are retained for analysis. Three such factors emerged from the data, and the factor loadings (after rotation) are shown for each item in Table 5.

For further clarification of the manner in which the items tended to fall into these three factors, Table 6 demonstrates graphically with the rotated data the relatively strong manner in which the questionnaire items aggregate into identifiable clusters. This table represents three visual-spatial displays of the position of each item in relation to two of the three factors. The top portion of Table 6, for example, displays Factor 1 on the vertical axis and Factor 2 on the horizontal axis. The manner in which items cluster relative to these factors is thereby represented. With three factors resulting, the
TABLE 6: TAPE EVALUATION
FACTOR ANALYSIS OF QUESTIONNAIRE ITEMS

Factor 1

Factor 2

Factor 3

Factor 2

Factor 3
### TABLE 7
QUESTIONNAIRE ITEMS LOADING INTO EACH FACTOR

<table>
<thead>
<tr>
<th>Loading Coefficient</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor 1: General, Content</strong>*</td>
<td></td>
</tr>
<tr>
<td>.691</td>
<td>1. The ideas in this tape were clearly and efficiently presented.</td>
</tr>
<tr>
<td>.771</td>
<td>6. I would recommend this tape to a colleague interested in this topic.</td>
</tr>
<tr>
<td>.839</td>
<td>7. Useful information was given which is applicable in a clinical situation.</td>
</tr>
<tr>
<td>.698</td>
<td>9. The presentation of this topic was well organized; good introduction, presentation and summary.</td>
</tr>
<tr>
<td>.839</td>
<td>10. The content of this tape was what I expected from reading the title.</td>
</tr>
<tr>
<td>.760</td>
<td>12. In generaly, I would say that this is a good tape.</td>
</tr>
<tr>
<td>.724</td>
<td>13. From what I know of the subject, the content of this tape is generally accurate.</td>
</tr>
<tr>
<td>.775</td>
<td>16. This tape is useful for general (continuing) education of the intended listener.</td>
</tr>
<tr>
<td>.714</td>
<td>17. The content of this tape seems to me to be out of date.</td>
</tr>
<tr>
<td><strong>Factor 2: Speaker Performance</strong> **</td>
<td>**</td>
</tr>
<tr>
<td>.810</td>
<td>2. I felt the speaker was a real expert talking to me personally.</td>
</tr>
<tr>
<td>.735</td>
<td>4. The speaker is not an authority on the subject. He/she was just reading.</td>
</tr>
<tr>
<td>.720</td>
<td>5. The speaker had a good voice, one that was pleasing to listen to.</td>
</tr>
<tr>
<td>.778</td>
<td>8. The speaker did not read the script well. He sounded monotonous.</td>
</tr>
<tr>
<td>.809</td>
<td>15. I thought the speaker was interested in his subject.</td>
</tr>
</tbody>
</table>
TABLE 7, Continued
QUESTIONNAIRE ITEMS LOADING INTO EACH FACTOR

<table>
<thead>
<tr>
<th>Loading Coefficient</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>.879</td>
<td>3. The tape was very clear; there was no noise or distortion.</td>
</tr>
<tr>
<td>.882</td>
<td>14. There was a lot of noise on the phone line, making it hard to understand.</td>
</tr>
</tbody>
</table>

Factor 3: *Technical Quality***

visualization of the relationships is relatively easy. One should imagine a sphere, with each rotated factor falling along one of the three dimensions or planes intersecting the sphere. Note, by the way, that item 11 does not fall into any one of the clusters very strongly. A review of Table 5 will show that this item is the only one with a communality level of less than .500, an indication of an item weak in discriminating ability among the factor clusters. For this reason, the item was not used any further. All other items were retained for subsequent analysis.

Table 7 presents a listing of each questionnaire item by factor cluster, and it shows the loading coefficient for each within that factor. Following the factor analysis, a factor label was assigned. Factor 1, which has an Eigenvalue of 8.73 and which in itself represents over 50 percent of the total variability in the questionnaire, is labeled *General, Content*. These items related to perceived general value of the tape and the more specific dimensions of content organization and validity. Factor 2 clearly deals with measures of the capabilities and credibility of the person speaking, and is called *Speaker Performance*. The third factor, composed of the two remaining items, deals specifically with and is labelled *Technical Quality*.

In addition to determining the factor dimensions and identifying the items belonging to each, the correlational relationship between the factors is of interest and is shown in Table 8. The main item of practical significance is the .665 Pearson correlation of the *General, Content* factor with the *Speaker Performance* factor. This indicates a
TABLE 8
PEARSON CORRELATION OF QUESTIONNAIRE FACTORS

<table>
<thead>
<tr>
<th>Factor pairing</th>
<th>correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1 by Factor 2</td>
<td>.665*</td>
</tr>
<tr>
<td>Factor 1 by Factor 3</td>
<td>.318*</td>
</tr>
<tr>
<td>Factor 2 by Factor 3</td>
<td>.233*</td>
</tr>
</tbody>
</table>

*p < .01

moderately high degree of association between the evaluator's perceptions of the general value of each tape and its contents with the performance and credibility of the speaker.

The results which are the major interest of this investigation are reported in Tables 9 and 10. These data matrices represent the degree of relationship between the call-back behavior reflected in the z-scores of the two groups of four tapes selected for evaluation and the scores obtained by each tape on the questionnaire. Questionnaire results were analyzed by the Pearson correlation statistic in relation to the call-back z-scores (Table 10). This correlational assessment was made on the basis of the total score for each factor and for all factors together.

The correlation test results are such that the acceptance of the second hypothesis is obtained on every parameter, with the strongest relationship being described between the Speaker Performance questionnaire factor (both tape groups taken together) and the call-back behavior of Dial-A-Tape users:

There will be a positive correlation between deviation from the average rate of call-back behavior for first-time callers of each tape and a rating of quality factors and perceived usefulness of that tape by a panel of representative health professionals.

The relationship between call-back behavior and the Speaker Performance factor is particularly strong for Tape Group 1. The Technical Quality factor also shows a fairly strong correlation with the call-back z-score for Tape Group 1, even stronger than the General, Content factor/z-score relationship.
### TABLE 9
TAPE SCORES: CALL-BACK & QUESTIONNAIRE RESULTS

<table>
<thead>
<tr>
<th>Tape Number</th>
<th>Call-back Z-score</th>
<th>Call-back Questionnaire</th>
<th>Questionnaire Factor 1</th>
<th>Questionnaire Factor 2</th>
<th>Questionnaire Factor 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tape Group 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>2.636</td>
<td>0.445</td>
<td>0.423</td>
<td>0.289</td>
<td>1.156</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>0.881</td>
<td>0.390</td>
<td>0.198</td>
<td>0.322</td>
<td>0.910</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-0.997</td>
<td>-1.277</td>
<td>-0.943</td>
<td>-0.298</td>
<td>-2.518</td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>-2.128</td>
<td>0.047</td>
<td>-0.555</td>
<td>0.081</td>
<td>-0.427</td>
<td></td>
</tr>
<tr>
<td>Tape Group 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>255</td>
<td>2.663</td>
<td>0.437</td>
<td>0.169</td>
<td>0.180</td>
<td>0.786</td>
<td></td>
</tr>
<tr>
<td>406</td>
<td>1.031</td>
<td>0.140</td>
<td>0.695</td>
<td>0.392</td>
<td>1.227</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>-0.913</td>
<td>-0.280</td>
<td>0.381</td>
<td>-1.319</td>
<td>-1.219</td>
<td></td>
</tr>
<tr>
<td>208</td>
<td>-2.740</td>
<td>0.175</td>
<td>-0.147</td>
<td>0.247</td>
<td>0.275</td>
<td></td>
</tr>
</tbody>
</table>

*all are standard scores

The second tape group has a generally weaker, though still meaningful relationship across all variables described, with the third factor dropping to a very weak correlation level. Tape Group 2 also shows a reversal of ranking of factors related to the call-back z-score, with the General, Content factor being slightly stronger than Speaker Performance.

Looking at all eight tapes together, the Speaker Performance factor emerges as the strongest predictor of call-back behavior (.587) with Technical Quality rather weak (.315).
### TABLE 10

**PEArson Correlation**:  
**Call-back Scores with Questionnaire Scores**

<table>
<thead>
<tr>
<th>Call-back Z-score</th>
<th>By Questionnaire</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1</td>
<td>Factor 2</td>
<td>Factor 3</td>
<td>Total</td>
</tr>
<tr>
<td>Tape Group 1</td>
<td>.539</td>
<td>.871</td>
<td>.629</td>
<td>.694</td>
</tr>
<tr>
<td>Tape Group 2</td>
<td>.517</td>
<td>.486</td>
<td>.251</td>
<td>.495</td>
</tr>
<tr>
<td>1 &amp; 2 Together</td>
<td>.456</td>
<td>.587</td>
<td>.315</td>
<td>.573</td>
</tr>
</tbody>
</table>

*Note: The normal approach to determining significance level does not seem to apply in this comparison. The data is in fact reduced to four cases, and thus technically involves but three degrees of freedom (or when analyzing tape groups 1 and 2 together, seven degrees of freedom.) But the results are really based not on just four or eight cases, but on the judgements or decisions of 46 people in group 1 and 37 people in group 2 regarding the questionnaire, and 80 other people in group 1 and 79 others still in group 2 regarding call-back behavior. There is, of course, a big difference in the degree of significance which can be reported depending on how one assesses this problem. For $df = 3$, $p = .05$ when the correlation coefficient is .878; for $df = 7$, $p = .05$ when the correlation coefficient reaches .666. But for $df = 36$, $p = .05$ when the correlation coefficient is as low as .320. While the latter numbers may not be entirely appropriate either, it is the investigator's contention that the evidence of the correlation between the two kinds of behavior reported, involving this number of people, is much stronger than the usual statistical approach to a four case or eight case analysis would indicate.*

---

**Questionnaire Interaction Analysis**

As a check on the validity of the questionnaire results, several statistical evaluations were made which assess the degree to which the evaluation procedure itself may have distorted the results. The following tables of data reflect no significant interactions except along dimensions the questionnaire was designed to measure.

Table 11 presents an analysis of the relationship between the number of questionnaires returned for each tape and the position in the order of hearing by the respondents. (The order of hearing was systematically ordered for the entire sample; all possible orders were equally represented among the total number of questionnaire
TABLE 11
INTERACTION ANALYSIS: NUMBER OF QUESTIONNAIRES
RETURNED BY ORDER OF HEARING

<table>
<thead>
<tr>
<th>Tape Number</th>
<th>First</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>11</td>
<td>8</td>
<td>10</td>
<td>17</td>
</tr>
<tr>
<td>33</td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>49</td>
<td>13</td>
<td>12</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>55</td>
<td>9</td>
<td>13</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>88</td>
<td>13</td>
<td>13</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>208</td>
<td>10</td>
<td>12</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>255</td>
<td>7</td>
<td>10</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>406</td>
<td>10</td>
<td>6</td>
<td>11</td>
<td>10</td>
</tr>
</tbody>
</table>

Chi-square = 18.03, with 21 degrees of freedom
Significance = .647

packets which were mailed out.) The relationship of these two variables, as measured by
the Chi-square statistic, was found to be insignificant.

An analysis of variance was made of the relationship between the item scores for
each questionnaire factor and the two tape groups, each tape number, the order of
hearing, the professional status of the respondents, and the incidence of free comments
on the questionnaires. Variance within the individual tape variable was found to be
significant, which is of course desirable. There is also significant variance between the
scoring for Tape Group 1 and Tape Group 2. This can be predicted by observing the
difference in scores for each group in Table 9; there is a greater spread between the lowest
and highest total scores and factor scores in Tape Group 1 than in Tape Group 2. All
other parameters were found to be insignificant, indicating support for the validity of the
questionnaire instrument and administration procedure.

The free comments section had been added to the questionnaire to allow for the
expression of dimensions of evaluation not anticipated in the questionnaire items. A
review of the comments made by respondents provided no points of measure different
from the items, so no further analysis was carried out on them.

In summary, the results of the investigation support the acceptance of both
hypotheses, demonstrate the manner in which health professionals perceive the
dimensions of the tapes among three major factors, and provide support for meaningful
TABLE 12
INTERACTION ANALYSIS: ANOVA OF QUESTIONNAIRE ADMINISTRATION VARIABLES

<table>
<thead>
<tr>
<th>Administration variable</th>
<th>Values of F</th>
<th>Degrees of freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1</td>
<td>Factor 2</td>
</tr>
<tr>
<td>Tape group</td>
<td>3.86*</td>
<td>21.07**</td>
</tr>
<tr>
<td>Tape number</td>
<td>20.70**</td>
<td>17.20**</td>
</tr>
<tr>
<td>Order of hearing</td>
<td>1.55</td>
<td>2.04</td>
</tr>
<tr>
<td>Professional status</td>
<td>0.65</td>
<td>0.12</td>
</tr>
<tr>
<td>Incidence of comments</td>
<td>0.02</td>
<td>1.67</td>
</tr>
</tbody>
</table>

*p < .05

**p < .01

conclusions about significant relationships between those factors and the actual behavior of callers toward the Dial-A-Tape system. These inferences include the observation that there is a moderately strong relationship between the panel of judges' perceptions of the general value of a tape, its content and structure, and their perceptions of the performance of the speaker. Also, the relationship between the judges' collective perceptions of speaker performance and the call-back behavior of the entire Dial-A-Tape user population is the strongest among the three factors found to exist in the questionnaire results.
CHAPTER IV
CONCLUSIONS, DISCUSSION, AND RECOMMENDATIONS

Introduction

This investigation was concerned with the development and testing of an effective method of evaluating individual tapes in a dial-access audiotape information system. Of particular interest was the evaluation of tapes from a communications effectiveness standpoint (rather than from the appropriateness of topic, which had been the focus in other dial-access evaluation studies.)

The study consisted of an evaluation design applied to the Arizona Regional Medical Program’s “Dial-A-Tape” service. This is a medical information service for physicians and nurses throughout the state of Arizona. More specifically, the objectives of the investigation were:

1. To measure the relationship between specific, first-selected tapes and the incidence of further calls to the tape library, and to use this data as one means of rating those tapes as an evaluation measure;
2. To develop a comprehensive, minimally reactive rating instrument for a subjective evaluation of individual tapes based on a number of communication variables;
3. To obtain from a representative group of health professionals their evaluation, using the rating instrument, of a sample of those tapes evaluated in step 1;
4. To perform a factor analysis on the rating instrument to determine the significant factors operating in this kind of evaluation of the tapes by health professionals, and the relationship among those factors; and
5. To determine the correlation between the call-back behavior rating and the questionnaire evaluation data for those tapes reviewed in step 3, both in
terms of overall rating and on each of the several factors found to exist in the questionnaire.

The hypotheses generated to fulfill the research objectives, and which arise out of the literature review are that:

1. There will be found in the records of all calls to the Arizona Dial-A-Tape service significant deviations from the overall mean rate of call-back behavior of first-time callers for specific tapes, with some tapes obtaining higher and others lower rates of call-back than can be attributed to chance.

and

2. There will be a positive correlation between deviation from the average rate of call-back behavior for first-time callers of each tape and a rating of quality factors and perceived usefulness of that tape by a panel of representative health professionals.

A review of the literature related to the evaluation of similar dial-access systems, to the nature of communications effectiveness dimensions which apply to this method and medium, and to certain problems generally occurring in social science research was conducted. The use of archival data related to the dial-access service was explored for its possible evaluation value; a related effort to develop and apply a survey instrument which would provide multi-dimensional evaluative data on individual tapes was made. The correlation of the results of the archival records measuring scheme and of the minimally reactive survey questionnaire provides the significant information resulting from this investigation.

On the theoretical assumption that the degree to which individuals return to use an information resource library is a result of their perception of the value of their first experience with that library, the variation in the rate of call-back for specific first-selected tapes is assumed to be a measure of those perceptions. The actual variation in this measure was determined as part of this investigation, and the results are that provided, therefore, a test of the first hypothesis.

To assess the call-back rating of tapes, all the calls to Dial-A-Tape for the first two and a half years of the existence of the service were recorded and ordered so that all the calls made by any given individual were listed together in chronological order. By means
of computer analysis, the first-called tape for every caller was tabulated. In addition, for each first-call a determination was made as to whether the caller ever called a second time to hear any tape in the dial-access library. The incidence of call-back following first-calls for each tape was accrued in this analysis. The rate of call-back for each first-selected tape was determined by this procedure; also resulting was the overall rate of call-back.

A survey questionnaire was developed which focused on those variables over which the producer of a tape has control, once the topic and general level of sophistication are determined. Called "communication parameters" in this study, the items in the questionnaire dealt with speaker performance, content organization, perceived usefulness in clinical application, accuracy of title, technical quality, and so forth.

After an informal review of the instrument's validity and a test-retest application to determine reliability, a final draft was administered to a state-wide randomly selected panel of physicians and nurses. Each was asked to call four designated tapes in the Dial-A-Tape library and fill out a copy of the questionnaire on each.

Two different groups of four tapes were used, with the specific tapes being selected from key points along the distribution of scores of call-back rating. In each tape group there was one from among the highest and one from the lowest scores (those tapes which met the significance criteria for the first hypothesis.) In addition, the group included a tape with a call-back z-score of approximately +1.0 and one with a z-score of about —1.0. A factor analysis was performed on the returned questionnaires, and the scores for each tape on each factor, as well as the total score for each tape, was correlated with the call-back z-scores for the tapes.

Seven of the 88 tapes eliciting ten or more first-calls accrued call-back rating z-scores which indicated significant deviation from the mean rate at a level of significance exceeding .05; (in addition, six more were significant at the .10 level.) Thus, the first hypothesis was accepted.

One hundred seventy-five of the 17-item questionnaires were sent out to 100 physicians and 75 nurses. A total of 83 were returned, 34 from the physicians and 49 from the nurses.

The factor analysis of the questionnaire resulted in three clearly defined factors relating to: 1) general evaluation and specific content considerations, 2) performance of the speaker, and 3) technical quality. A highly significant .665 correlation between the first two factors was found to exist. A much lower, but statistically significant correlation was found between the third factor and each of the first two.
To varying degrees a positive correlation was obtained between the call-back z-scores and the overall scores and each of the factor scores on the survey questionnaire. Due to the nature of the data handling and study design, the usual test of statistical significance does not realistically apply to these correlation coefficients; the degree to which chance operated in the results of this portion of the investigation is a matter of less than precise judgement, based on the manner in which the scores were attained for the 8 tapes involved. But essentially, the second hypothesis also was accepted.

Conclusions

The following conclusions are presented based on the results reported above:

1. Given a system of logging the name, location, and professional status of each caller, the logs from a dial-access library such as Dial-A-Tape can be the source of evaluation information on individual tapes within the dial-access library. This information is found in the call-back rating for each first-selected tape.

2. Such a measure is possible, however, only where a sufficiently large quantity of calls are made over a long period of time. Furthermore, such a measure is only possible when the identification of individual callers is accurate, or at least consistent.

3. Physicians and nurses tend to perceive the value of the tapes evaluated along three distinct dimensions: general value, heavily related to content accuracy, organization, clinical application, etc.; performance of the speaker; and technical quality. Though the factors are distinct enough, there is a high correlation between the first two dimensions, perceptions of the content and of the speaker. This correlation has practical significance to the producer of this kind of tape.

4. The correlations between the call-back data and the questionnaire results suggest that there is a relationship between call-back behavior and all three questionnaire factors. The strongest relationships are between call-back
behavior and users' perceptions of the structure and usefulness of the tape content and also of the performance of the speaker, with the speaker factor showing the strongest correlation of all.

5. Given the significant correlation between professional judges' perceptions of the general value of the tape and the performance of the speaker, and given that the highest correlation between call-back behavior and the questionnaire factors relates to speaker performance, it is concluded from these results that the user's perceptions of the speaker on these dial-access tapes has an influence on how the user values the dial-access library and the extent to which he returns to use it again.

6. Also, because the data demonstrate at least a moderate degree of relationship between the call-back behavior and the questionnaire results, it is concluded that the evaluation instrument designed for this study will have valid application in the formative evaluation of new tapes for this and perhaps other medically related dial-access audiotape information systems.

Discussion

On review of the call-back data relating to the first hypothesis, one might conclude that the normal properties of distribution are operating, but little else. Indeed, if at least several of the 88 tapes' scores did not support the hypothesized contention by approaching the extreme ends of a normal distribution, it would perhaps be even more significant. However, the investigation into this hypothesis does draw attention to the fact that a norm-referenced procedure for ranking tapes, based on behavior, is available through analysis of the dial-access system logs.

The extent to which practical meaning may be ascribed to this distribution was provided in the correlation of the questionnaire results. That evaluative data, obtained under the most representative yet controlled conditions deemed practical, yielded what appear to be highly valid, or at least minimally reactive, ranking assessments. Though the number of tapes evaluated by questionnaire was relatively small, the factor analyses and
the correlation results suggest that something other than random distribution (or an
element not measured) is operating on callers' behavior toward the use of Dial-A-Tape;
there is demonstrated a relationship between call-back behavior and perceptions of
quality. This investigator is cautiously assuming, in addition, that this is logically and by
definition a causal relationship.

An objective of the investigation was to design a viable and practical means of
assessing individual tapes through the use of archival data, and thus to avoid the pitfalls of
the more usual evaluation approach. It was assumed that once established as a viable
technique, doing this on a frequent or on-going basis would require the use of a
computer. While the conclusions suggest that this is indeed possible, the amount of error
found in the logs when manually reviewing all of the calling data indicates that such a
monitoring scheme could not be programmed into a computer without serious and costly
problems. To organize and review the log data manually was extraordinarily time
consuming; the cost for this kind of effort on an on-going basis would be prohibitively
high.

The technique used would lend itself to successful and cost-effective evaluation in
any system where a quick and accurate means of user identification could be applied.
While not yet practical, one can readily foresee the time when, by means of the total
wired communications systems now being envisioned for whole communities, automatic,
digital identification of callers would be routinely accomplished for all kinds of
information transactions. The great opportunity for error extant in the use of inconsistently spelled names on hand-recorded logs would then no longer be a limiting
factor. It is concluded by the investigator that computer processing and automatic
evaluation of user behavior toward an information system, based on the procedures and
premises described in this investigation, would then have valid, significant application in
the management of such a service.

In spite of the data quality and quantity limitations in this study, several comments
can be made as a result of this investigation relating to other dial-access evaluation efforts.
Berglund’s hunch that some tapes might be generating discontent with the Minnesota
dial-access service [1969] is supported in these results; poorly rated tapes do seem to “turn
people off" from using the system again in a most specific way. The mistake of assuming that frequently used tapes are highly valued is also demonstrated here. Tape number 13, "Marriage on the Rocks", has been reported among the "top ten" tapes by several different dial-access services using the Wisconsin-produced tapes. Yet, it was valued by the professional judges in this investigation as the worst tape of the eight reviewed.

The physician community has not appeared to take very seriously the need for involving competent, trained communications experts in the development of dial-access information services. As a result, evaluative concern focused instead on the appropriateness and potential usefulness of certain topics, the number of messages in the dial-access library, the cost of accessing it and so on. However, human beings are involved; and even physicians react like most other people with biases and perception of believability, intelligibility, and effectiveness.

A number of studies in the speech communication area were reviewed in Chapter 1. While these dealt primarily with persuasive speeches, the similarity in results between those studies and this one is impressive. Knowler [1936] and Heinberg [1963] made a strong case for the essentially inseparable relationship of speaker and content to the effectiveness of the communication. This relationship is strongly indicated here, too, with the high degree of correlation between the speaker performance and the general—content factors in the questionnaire results. One can find support as well in these results for Andersen and Clevenger’s claim [1963] as to the universal relationship between elements of ethos and the impact of the message.

Generally, if the listener of a taped message in the Dial-A-Tape library makes judgments as to the personality and character of the speaker, as suggested by Cantril and Allport [1935], if his reaction to the content of the tape is positively correlated to the perceived personality of the speaker through the latter’s voice and delivery [Knowler, 1936], and if his further utilization of such an informational library is correlated to his perceptions of the speaker as well as the content as suggested in the results of this study, the conclusion is clear: as with the general forms of public speaking, attention must be paid by the communicator of health information via Dial-A-Tape and similar systems to factors of speaker performance. These include delivery, voice quality, and rate of speaking. The need for a tape to attract and hold attention and to communicate in the aural dimension in
competition with distractions, fatigue, impatience with perceived incompetence or disinterest has been demonstrated; this is in addition to the requirement now readily recognized for accuracy and inherent information value in the taped message.

Recommendations for Further Research

To more fully deal with the questions raised in this investigation, the following suggestions are made for further research:

1. A replication of this study should be made in a similar dial-access information library system where the data base is larger and where the identification of the callers can be accurately determined. The replication would also ideally include the questionnaire evaluation of a much larger proportion of the tapes.

2. A second, more limited suggestion, is the amplification of this study by having evaluated a number of additional tapes chosen from key points along the call-back rating distribution reported here in order to further confirm (or deny) the relationship between call-back ratings and professional judgement.

3. It is further recommended that in-depth interviews be conducted with first-time callers about one month after their first call to determine their assessment of the first-tape selected. A key dimension would be a comparison of those who had called back in the meantime and those who had not, (thus the need for some delay between the first-call and the interview.)

4. Another approach to the more accurate determination of the effect which various communications parameters might have on the perceived value of such tapes would be to systematically vary one or more such dimensions in a particular set of tapes which are listed as part of the dial-access library and then to monitor call-back behavior and/or apply the questionnaire process. A special opportunity might exist in a relatively new dial-access library, where a high proportion of first-calls are being placed. Different versions of the tapes being manipulated could be alternately played during different time
periods each day, or on alternate days or weeks. Or they could simply be rotated every time the tape number is accessed. Careful notation as to which caller heard which versions of these tapes would, of course, be essential to a successful study. Also essential would be the accumulation, once again, of a large enough data base over a long enough period of time to obtain statistically meaningful information.

*Recommendation for Design and Implementation of a Dial-Access System*

It is recommended that the manager of a new dial-access medical information system set out a list of guidelines, or minimally acceptable criteria, based on the well established principles of effective public speaking found in a wide variety of sources. These criteria should apply to the development of every tape in the library. Further, it is recommended that the evaluation instrument developed in this study, or an expanded version of it, be utilized in the formative evaluation of every tape. This should be applied using the collective judgements of a sample of potential users of the service, not just professional colleagues who happen to be available.

It is strongly suggested, too, that persons trained in presenting information via this medium should be employed to actually record the tapes. There is some risk in maintaining credibility with even a professional narrator, however, if he is unfamiliar with the medical area even though well trained otherwise. The ideal situation would be to find people trained to some degree in the health sciences who are also very effective oral communicators.

Given these personnel resources, an effective professional alliance may be created between the health professional content specialist and the skilled medical communicator. The application, then, of the kind of evaluation instrument and monitoring techniques suggested in this investigation can be made on a continuing basis. A combination of elements, including the commitment to meet the continuing education needs for health professionals outlined earlier, the tremendous wealth of new information to be communicated, the audio communication technology available in the tape recorder and the telephone, and the practical application of communications research data such as those reported herein should provide a powerful and effective educational resource in the health sciences.
Effective continuing education services to the health sciences community are important. Dial-accessed audiotapes of health information have been employed, and more effective means of evaluation are needed. The use of actual behavior of users of such an information service, as recorded in the logs of calls made, was perceived as a valuable means of obtaining evaluation of individual tapes in the Arizona Dial-A-Tape library.

Variability in the rate at which first-time users of such a service call back again was found to exist; it is suggested that this variability is a means to discriminate among the tapes the perceived quality of those tapes. A questionnaire pertaining to communications-related concerns was designed and administered to a state-wide sample of health professionals who were asked to rate a few tapes selected from key positions among the distribution of call-back data.

The resulting factor analysis and correlation measures suggest that call-back behavior is in fact related to perceptions of the value of first-called tapes, and that these perceptions operate along three dimensions: general value, and in particular, content-related matters; performance of the speaker; and technical quality. All factors appear to interact at a meaningful level, and overall, speaker performance appears to correlate most highly with call-back behavior.

If one accepts the conclusion that the call-back behavior supports the results of the questionnaire, and that the correlation between call-back behavior and the first two factors in particular is meaningful, the implications for design and production of future tapes are clear. Successful taped messages must be prepared in a well organized and efficient manner, designed for hearing over the telephone. They must be presented by individuals who sound interested, and who seem to be highly credible, articulate people. The use of health professionals whose primary qualification is an interest in the project and a willingness to donate effort toward the production of tapes is not sufficient. The manager of a dial-access system must, according to the results, involve persons skilled in effective communications to a greater degree than has been the case in the development of those tapes used in the Arizona Dial-A-Tape service.
The results of this investigation appear to confirm and substantiate the findings of earlier studies regarding communications effectiveness, as well as support the validity of the archival record evaluation technique. However, this study dealt with a small quantity of data and with but a single audiotape information service. Additional research using a broader data base and involving the evaluation of a larger number of tapes is suggested to substantiate these findings. Also, the variables pertaining to the use of professional communicators in such taped messages should be more thoroughly investigated for specific implications in continuing education information services.
APPENDIX A

DIAL-A-TAPE CALL LOG SHEET
<table>
<thead>
<tr>
<th>Day</th>
<th>Mo.</th>
<th>Yr.</th>
<th>Hour</th>
<th>Requestor's name</th>
<th>Prof. Stat.</th>
<th>Requestor's town</th>
<th>Tape No.</th>
<th>Other info requested Yes</th>
<th>No</th>
<th>Telephone number</th>
<th>Have you listened to our tapes before?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B

PRELIMINARY DRAFT OF THE QUESTIONNAIRE
DIAL–A–TAPE EVALUATION QUESTIONNAIRE

Number and title of tape: ____________________________________________________________

Your professional status: _______________________________________________________________________________

Please select any four digit code number which you can remember, such as the last four digits of your phone number, social security number, home address, etc. Put this code number on the line in the upper right-hand corner of this page. Use the same code number on all copies of this form which you fill out.

First call "Dial–A–Tape" (884-3351) and listen to the tape. Then, select the response which most closely represents your first impressions regarding each item and mark it with a check. (Please be sure to answer every item.)

1. In general, I would say that this is a good tape.

   __________ / __________ / __________ / __________ / __________
   Strongly Disagree Disagree Neutral Agree Strongly Agree

2. From what I know of the subject, the content of this tape is generally accurate.

   __________ / __________ / __________ / __________ / __________
   Strongly Disagree Disagree Neutral Agree Strongly Agree

3. The title of this tape is misleading.

   __________ / __________ / __________ / __________ / __________
   Strongly Disagree Disagree Neutral Agree Strongly Agree

4. I felt the speaker was a real expert talking to me personally.

   __________ / __________ / __________ / __________ / __________
   Strongly Disagree Disagree Neutral Agree Strongly Agree

5. There was not enough summarizing of main points.

   __________ / __________ / __________ / __________ / __________
   Strongly Disagree Disagree Neutral Agree Strongly Agree

6. The content of the tape was what I expected from reading the title.

   __________ / __________ / __________ / __________ / __________
   Strongly Disagree Disagree Neutral Agree Strongly Agree

7. I felt the speaker was unsure of himself, lacked confidence.

   __________ / __________ / __________ / __________ / __________
   Strongly Disagree Disagree Neutral Agree Strongly Agree
8. The presentation of this topic was well organized: good introduction, presentation and summary.

9. Useful information was given which is applicable in a clinical situation.

10. The speaker talked at just the right rate of speed (not too fast nor too slow).

11. I think the tape contained some erroneous information.

12. The ideas in this tape were clearly and efficiently presented.

13. The speaker sounded muffled; the recording was technically poor.

14. The content of this tape seems to me to be out of date.

15. He (she) spoke so slowly my mind began to wander.

16. This tape is useful for general (continuing) education of the intended listener.

17. The tape was very clear; there was no noise or distortion.
18. The speaker is not an authority on the subject. He was just reading.

19. The basic ideas or thoughts were presented so fast I got lost.

20. I thought the speaker was interested in his subject.

21. There was a lot of noise on the phone line, making it hard to understand.

22. The speaker had a good voice, one that was pleasing to listen to.

23. No specific, useful steps were given to manage a patient with the problem described.

24. The speaker did not read the script well. He sounded monotonous.

25. I would recommend this tape to a colleague interested in this topic.

Thank you very much for your assistance!
APPENDIX C

QUESTIONNAIRE AS ADMINISTERED TO

SAMPLE OF HEALTH PROFESSIONALS
Dial—A—Tape Evaluation Questionnaire

Number and title of tape:  

Date of call: ___________ Time called: _______ Your professional status (MD, RN, etc.): _____________________

Please select any four digit code number, such as the last four digits of your phone number, social security number, home address, etc. Put this code number on the line in the upper right-hand corner of this page. Use the same code number on all copies of this form which you fill out.

Next, call “Dial-A-Tape” and request to hear the tape listed above. (In Maricopa County, call 252-5581. All others call 884-3351 in Tucson. Call collect where necessary.) Then fill out this evaluation form, selecting the response which most closely represents your first impressions regarding each item and mark it with a check. (Please be sure to answer every item.)

1. The ideas in this tape were clearly and efficiently presented. 
   Strongly Disagree Neutral Agree Strongly

2. I felt the speaker was a real expert talking to me personally. 
   Strongly Disagree Neutral Agree Strongly

3. The tape was very clear; there was no noise or distortion. 
   Strongly Disagree Neutral Agree Strongly

4. The speaker is not an authority on the subject. He/she was just reading. 
   Strongly Disagree Neutral Agree Strongly

5. The speaker had a good voice, one that was pleasing to listen to. 
   Strongly Disagree Neutral Agree Strongly

6. I would recommend this tape to a colleague interested in this topic. 
   Strongly Disagree Neutral Agree Strongly

7. Useful information was given which is applicable in a clinical situation. 
   Strongly Disagree Neutral Agree Strongly
<table>
<thead>
<tr>
<th></th>
<th>The speaker did not read the script well. He sounded monotonous.</th>
<th>Strongly Disagree / Disagree / Neutral / Agree / Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>The presentation of this topic was well organized; good introduction, presentation and summary.</td>
<td>Strongly Disagree / Disagree / Neutral / Agree / Strongly Agree</td>
</tr>
<tr>
<td>10</td>
<td>The content of this tape was what I expected from reading the title.</td>
<td>Strongly Disagree / Disagree / Neutral / Agree / Strongly Agree</td>
</tr>
<tr>
<td>11</td>
<td>The speaker talked at just the right rate of speed (not too fast nor too slow.)</td>
<td>Strongly Disagree / Disagree / Neutral / Agree / Strongly Agree</td>
</tr>
<tr>
<td>12</td>
<td>In general, I would say that this is a good tape.</td>
<td>Strongly Disagree / Disagree / Neutral / Agree / Strongly Agree</td>
</tr>
<tr>
<td>13</td>
<td>From what I know of the subject, the content of this tape is generally accurate.</td>
<td>Strongly Disagree / Disagree / Neutral / Agree / Strongly Agree</td>
</tr>
<tr>
<td>14</td>
<td>There was a lot of noise on the phone line, making it hard to understand.</td>
<td>Strongly Disagree / Disagree / Neutral / Agree / Strongly Agree</td>
</tr>
<tr>
<td>15</td>
<td>I thought the speaker was interested in his subject.</td>
<td>Strongly Disagree / Disagree / Neutral / Agree / Strongly Agree</td>
</tr>
<tr>
<td>16</td>
<td>This tape is useful for general (continuing) education of the intended listener.</td>
<td>Strongly Disagree / Disagree / Neutral / Agree / Strongly Agree</td>
</tr>
<tr>
<td>17</td>
<td>The content of this tape seems to me to be out of date.</td>
<td>Strongly Disagree / Disagree / Neutral / Agree / Strongly Agree</td>
</tr>
</tbody>
</table>
Please write below any other comments you have about this tape.

Thanks very much for your assistance. Be sure to return questionnaires in the preaddressed envelope to Medical Audiovisual Services, Arizona Medical Center, Tucson, Arizona 85724.
APPENDIX D

"ADVANCE" LETTER SENT TO SAMPLE
I am conducting a study of the Dial-A-Tape medical information service which has been available to physicians and other health professionals here in Arizona for the past three years. We are gathering information which will assist in the design and improvement of such information systems.

Dial-A-Tape is an audiotape library service, provided through the Arizona Regional Medical Program in cooperation with the Arizona Medical Center at the University of Arizona. It consists of a library of several hundred 5- to 8-minute audiotape recordings on selected medical topics. All nurses and physicians registered in the state have received copies of the catalog at one time or another.

A tape is heard by calling the Dial-A-Tape service in either Tucson or Phoenix. The caller is asked for some brief identification data and for the number of the tape he wishes to hear. It is then played over the phone. The phone disconnects automatically at the end of the tape.

In a few days you will be receiving from me a request for your assistance in this study. I will appreciate it very much if you will take the time to participate (about 30 - 40 minutes altogether.) You, as one of a fairly small sample of health professionals in Arizona will be asked to call the Dial-A-Tape service and listen to each of four specified tapes. Following each tape, a brief questionnaire is to be completed. Then the questionnaires are to be returned to me in a postage-paid envelope.

Your participation and cooperation will be greatly appreciated and will assist us in developing information services for health professionals in Arizona and elsewhere in the country.

Thanks very much, in advance, for your help.

Cordially,

William T. Harrison
Assistant Director, Medical Audiovisual Services
APPENDIX E

COVER LETTER ACCOMPANYING QUESTIONNAIRE PACKETS
A few days ago I wrote and told you that you had been selected to participate in a research project designed to assess the Dial-A-Tape medical information service here in Arizona. We are developing research methodology to facilitate improvement in telephone audiotape information systems such as Dial-A-Tape, and your help is needed at this point.

Enclosed are four copies of an evaluation questionnaire, one for each of four tapes. The reference number and title of each tape is already typed at the top of each questionnaire. Please follow the instructions regarding code number and professional status (MD, RN, etc.) Then call and listen to the tape indicated. The phone number you are to call is printed on the questionnaire.

It is not necessary to listen to all four tapes in the same sitting. You may call whenever you like (Dial-A-Tape service is available 24 hours a day), but please complete all four within two days or so. In any case, it is important that you make your calls, fill in the questionnaires, and get them returned within the next week. It is also important that you listen to them in the sequence order indicated; a number written in the upper left-hand corner indicates which is to be listened to first, which second and so on.

The tapes you are asked to evaluate were chosen to correlate with other data we are gathering and do not relate necessarily to your particular professional status or specialty. To the degree possible, assume you are making judgements as a person for whom the content of the tape is intended; we are asking for your responses as a professional evaluator of the tape as well as a potential user of the information it contains.

Your cooperation and assistance in this research project will be greatly appreciated. Please do not lay this aside; the deadline for your replies is a very real one — the project will not be successful if replies are not returned within a week.

Thanks very much!

Cordially,

William T. Harrison
Assistant Director,
Medical Audiovisual Services
APPENDIX F

FOLLOW-UP LETTER SENT TO SAMPLE
A number of days ago you received a request to participate in a study of the Dial-A-Tape medical information service being provided to physicians and nurses in Arizona by the Arizona Regional Medical Program. We are seeking data which may be applied to the improvement of such information systems for medical personnel.

According to the record of calls (as of November 1) you have not yet had time to participate as requested. In case this has, by now, slipped your mind, could I again request your assistance? Because the capacity of the telephone answering equipment and personnel are limited, we have sent questionnaires to quite a small sample of physicians and nurses throughout Arizona. In order to obtain useful data with such a small sample, we must seek everyone's response; your efforts will count!

Remember, you can call collect if you are outside the metropolitan Phoenix or Tucson calling areas, and you need not call all four tapes at one sitting. (The phone number is printed on the questionnaires.) Your participation will be very much appreciated.

If you have indeed already made the calls requested, please accept my personal thanks for your assistance.

Cordially,

William T. Harrison
Assistant Director,
Medical Audiovisual Services
BIBLIOGRAPHY


Miller, George, and Harless, William G. "Instructional Technology and Continuing Medical Education," *To Improve Learning: A Report to the President and the Congress of the United States by the Commission on Instructional Technology.*


Oakley, Claron L. "Tape Recordings in Medical Education," California Medicine, 104 (June, 1966), pp. 534-7.


Ragatz, Roy. Personal communication to James W. Barry, Librarian at the University of Arizona Medical Center, January 20, 1972.


