RECOMMENDED STANDARDS FOR RADIO AND TELEVISION PROGRAM AUDIENCE SIZE MEASUREMENTS

ADVERTISING RESEARCH FOUNDATION
11 West 42nd Street, New York 36, N. Y.

RECOMMENDED STANDARDS FOR RADIO AND TELEVISION PROGRAM AUDIENCE SIZE MEASUREMENTS

WORKING COMMITTEE ON STANDARDS AND METHODS

G. MAXWELL ULE, Chairman
Kenyon & Eckhardt, Inc.

DONALD W. COYLE
American Broadcasting Company, Inc.

E. L. DECKINGER
The Biow Company, Inc.

JOHN MACDONALD
Committee Secretary
Advertising Research Foundation, Inc.

MEMBERS OF RADIO-TELEVISION RATINGS REVIEW COMMITTEE

E. L. DECKINGER, Chairman
Vice President, Director of Research
The Biow Company, Inc.

MRS. "TEDDY" ANDERSON
Manager of Radio-TV Research
Batten, Barton, Durstine & Osborn, Inc.

HUGH M. BEVILLE, JR.
Director of Research and Planning
National Broadcasting Company, Inc.

HARPER CARRAINE
Director of Research
CBS Radio Network

DONALD W. COYLE
Director of Research
American Broadcasting Company, Inc.

WALLACE T. DREW
Account Executive
Grey Advertising Agency, Inc.

JAY ELIASBERG
Coordinator of Special Projects
American Broadcasting Company, Inc.

NORMAN W. GLENN
Vice President
Doherty, Clifford, Steers & Shenfield, Inc.

ROBERT M. HOFFMAN
Director of Planning & Development
WOR, WOR-TV

GORDON A. HUGHES
Director, Market Analysis Department
General Mills, Inc.

OSCAR KATZ
Director of Research
CBS Television

HOWARD G. KUHN
Vice President, Manager Research Department
Compton Advertising, Inc.

WILLIAM J. MILLARD, JR.
Manager of Radio & TV Research
McCann-Erickson, Inc.

HAL MILLER
Manager of Media Analysis
Benton & Bowles, Inc.

RICHARD J. PUFF
Director of Research & Sales Planning
Mutual Broadcasting System, Inc.

BUD SHERAK
Account Research Director
Kenyon & Eckhardt, Inc.

CHARLES H. SMITH
Research Director
WCCO

G. MAXWELL ULE
Vice President, Director of Research
Kenyon & Eckhardt, Inc.

H. D. WOLFE
Director of Marketing Research
Colgate-Palmolive Company

ARF'S OFFICIALS AND DIRECTORS

OFFICIALS

MARION HARPER, JR.
Chairman*
President
McCann-Erickson, Inc.

EDGAR KOBAK, President*
Advertising Research Foundation, Inc.

A. W. LEHMAN
Managing Director
Advertising Research Foundation, Inc.

EDWARD F. HERRICK, Secretary
Advertising Research Foundation, Inc.

F. B. MANCHEE, Vice Chairman*
Executive Vice President
Batten, Barton, Durstine & Osborn, Inc.

E. P. SEYMOUR, Treasurer*
Vice President
The Crowell-Collier Publishing Company

D. B. LUCAS, Technical Director
Advertising Research Foundation, Inc.

GILBERT WEIL, General Counsel

DIRECTORS

FREDERIC R. GAMBLE
President
American Association of Advertising Agencies, Inc.

B. B. GEYER
President
Geyer Advertising, Inc.

WILLIAM A. HART
Director of Advertising
E. I. du Pont de Nemours & Company, Inc.

ROBERT E. KINTNER
President
American Broadcasting Company, Inc.

WALTER C. KURZ
Manager, General Advertising
Chicago Tribune

J. WARD MAURER
Director of Advertising
Wildroot Company, Inc.

FERGUS MEAD
Vice President
The Buchen Company

PAUL MONTGOMERY
Executive Vice President
Publications Division
McGraw-Hill Publishing Company, Inc.

ADRIAN MURPHY
President of CBS Radio
Columbia Broadcasting System, Inc.

VERNON C. MYERS
Publisher
Look Magazine

* Also a director.

NOTE

In accordance with regular Foundation procedures, this report was submitted to the Board of Directors for review and approval. It has been duly approved and its publication authorized by the ARF Board of Directors.

ARF AND ITS OBJECTIVES

The Advertising Research Foundation is a non-profit organization supported by advertisers, advertising agencies and advertising media. Its basic purpose is to further scientific practices and promote greater effectiveness in advertising and marketing through objective and impartial research.

Specifically, the Foundation assists advertising media in developing research data; supervises and validates media research surveys, and appraises media studies. Further, the ARF seeks to develop new research methods and techniques; to analyze existing techniques and define their proper application and limits of usefulness; and to establish research standards and criteria.

The ARF collects and disseminates advertising and marketing data for the benefit of subscribers. It is interested in exploring any research plan which might increase the industry's knowledge of how to make advertising more effective.

Contents

Foreword

Early in 1952, the ARF sent a questionnaire to its subscribers to determine their interest in 35 specific research projects. One of the projects favored by a great majority of subscribers was a study of radio and television rating methods.

With this mandate from subscribers, the ARF Board of Directors authorized the study and a committee for this project was organized almost immediately. The project became the first study of research methods to be undertaken by the ARF, and this report is the first of a series which ARF plans to issue on various phases of this subject.

GENESIS AND IMPORTANCE OF THE PROBLEM

The problem developed because different program audience size measurement services have often reported widely divergent measurements for the same broadcast. The reasons for these differences in estimates can be grouped into three categories:[1]

1. Method—such as elapsed time between broadcast exposure and the act of obtaining the program audience size measurement information; measurement of average instantaneous or total audience; and measurement of in-home audience or in- and out-of-home audience.
2. Sample—such as type of sample, sample size and area covered.
3. Procedures and Processing—such as handling of tabulations, reporting period and particular broadcast covered.

Since the Industry cannot continuously tolerate widely divergent measurements for the same broadcast, the Radio-TV Ratings Review Committee was assigned the responsibility of developing standards to produce more uniform measurements.

[1] *A Plan for the Evaluation of Audience Measurement Methods*, published by the National Association of Broadcasters, Washington, D. C., March 1951.

PRESENTATION OF RESULTS

This report presents what your Committee hopes will become long-term standards in the Industry for program audience size measurements. It also includes your Committee's appraisal of the practical potential of each of the currently practiced methods at the present state of technology in research. This is primarily the product of the Working Committee on Standards and Methods which is one of four subcommittees of the Radio-TV Ratings Review Committee.

This report is not a critique of any individual program audience size measurement service as such. It is, rather, a study of what each method can and cannot accomplish if it is operated up to the limit of its fullest practical potentialities, regardless of who the operator might be. Neither is any effort made to appraise the extent to which particular practitioners do or do not employ their chosen methods in such a manner as to derive maximum results from them.

In a field as dynamic as radio and television program audience size measurements, it is inevitable that there should be a good deal of doctrinal disagreement among the experts. It is the competition of such opposing theories in the arena of public discussion that gives science its greatest opportunity to develop and validate principles.

The contents of this report are founded on and reflect the best collective judgments of the members of your Committee. These judgments are based on their knowledge and experience.

THE INDUSTRY'S RESPONSIBILITY

Your Committee has established high standards as goals for program audience size measurements because no real achievement is ever made unless the objective is high. To set low standards would imply that the Industry has reached the end of its inventive genius. This thesis your Committee refuses to accept. The program audience size measurement practitioners have shown great ingenuity and the Industry can be proud of their achievements. It is in the spirit of helping to build even further on these achievements that your Committee sets its standards. At the same time, however, your Committee has not set the standards so high that they are unattainable or unrealistic.

We are confident that the practitioners will continue to develop im-

provements that will bring us ever closer to the common objective of better knowledge about audience sizes.

In setting high standards your Committee appreciates that, for the present, there will be a gap between objective and attainment. What is very desirable may, at least at our present level of knowledge, be unattainable at reasonable cost.

Your Committee cannot, of course, foresee future developments in technology, knowledge and ideas. It is always possible that new developments may make it necessary to revise these recommended standards.

In a free society like ours, supply and demand rule. Important improvements will be forthcoming only if the buyers of program audience size measurements recognize their need and demand them. Moreover, in the economist's sense of demand being "desire backed by purchasing power" this objective cannot be attained if it is not supported by an appreciation of the fact that if we *want* good research, we must be prepared to *pay* for good research. In any field, you get what you pay for; good work *cannot* be had unless someone is willing to pay its cost.

Technically inadequate program audience size measurements can do more harm than good, because they will frequently lead to unreliable estimates—and wrong decisions. It is toward this goal of achieving improved program audience size measurements that your Committee is wholly dedicated; it is toward this goal that your Committee enlists the earnest and active support of every reader.

HOW THIS REPORT WAS WRITTEN

To develop these recommended standards, the Working Committee on Standards and Methods held approximately one hundred meetings and conferences. All of the problems involved in radio and television program audience size measurements were fully discussed. Ten different program audience size measurement methods were reviewed and analyzed.

At various stages the Working Committee submitted drafts of the report to the full committee for discussion. After each meeting of the full committee the Working Committee revised the draft in accordance with the decisions reached. This final draft was submitted to the entire committee for approval for publication. They unanimously recommended to the ARF Board of Directors that the report be published.

In line with ARF's established procedure designed to obtain the maximum amount of information and to gather all shades of opinions regarding research techniques and practices, each radio and television rating practitioner was asked for comments and suggestions after having been given ample time to study a draft of this report. These comments were presented to the Working Committee in written or verbal form or both. All of them were carefully studied and reviewed. They aided materially in increasing the scope and accuracy of this report since a substantial number of revisions in the context of the report resulted from this interchange of ideas. ARF acknowledges with deep appreciation and thanks the help and cooperation received from: American Research Bureau, Inc.; C. E. Hooper, Inc.; A. C. Nielsen Company; The Pulse, Inc.; Trendex, Inc.; and Videodex, Inc.

After the report was rewritten, the final draft was again sent to all the practitioners for further comments. The Working Committee again carefully considered and studied the comments received. The second set of comments resulted in six changes in the report. Details regarding these changes are given on page 70.

There remain a number of important points on which the various practitioners and your Committee do not agree. However, ARF procedure in preparing reports of this nature provides for giving each interested practitioner a full opportunity to express his viewpoints because ARF believes that their considered opinions in any phase of advertising or marketing research are worthy of serious attention. Consequently, their comments concerning this report are given in full as addenda.

PURPOSE OF THIS REPORT

This report recognizes there can be a difference between the standards for radio and television program audience size measurements and what a given method can achieve. There can also be a difference between what a method can deliver (when practiced to its fullest practical potential) and what it does deliver in present practice.

The purpose of this report is twofold:

1. *To outline your Committee's standards for radio and television program audience size measurements.* Once this objective is achieved, the users of program audience size measurements will

have fulfilled their responsibility of stating specifically what they need. Thus, the services will have a guide for improving their own measurements.

2. *To state the potential of each of the current program audience size measurement methods when practiced to their fullest practical potential.* Once this objective is achieved, the Industry will know what each method can do in obtaining the specified kinds and quality of program audience size measurement information. This analysis of the potential of the various methods may help to indicate the areas in which further development would be most fruitful.

Summary

This report presents your Committee's standards for radio and television program audience size measurements and also evaluates how well each of the principal program audience size measurement methods can meet these standards, assuming an optimum compromise among theory, practice and cost.

STANDARDS FOR PROGRAM AUDIENCE SIZE MEASUREMENTS

The standards set forth in this report for program audience size measurements define:

1. The type and amount of information to be provided.
2. The degree of accuracy required.
3. The procedures to be followed.

METHODS APPRAISED

This study appraises ten different program audience size measurement methods on the basis of the standards established in this report. Seven of the methods are independent and unique systems of measuring radio and television broadcast audiences. Each of the other three methods is a combination of two of the seven fundamental methods.

POTENTIAL OF METHODS

Analysis indicates that there are strengths and weaknesses in each of the methods studied. None of the methods can satisfy all of the standards. This is your Committee's judgment of the various methods reviewed in terms of their theoretical promise under optimum implementation. It remains for future analyses to determine the degree to which current practices of these methods deliver on their full potential.

STATISTICAL CONSIDERATIONS

Good sampling and samples of adequate size are necessary for valid and reliable program audience size estimates. Your Committee recommends the type of sampling procedure which it feels is desirable and suggests acceptable sample sizes for local and national measurements. For a complete understanding of the references to sample size and the conditions under which they apply, it is necessary to read Chapter IV.

VALIDATION OF PROGRAM AUDIENCE SIZE MEASUREMENT DATA

Two sets of criteria are recommended for the validation of program audience size measurements. One is the recently published ARF *Criteria for Marketing and Advertising Research*. The other is related to adequate record-keeping by each practitioner to permit further analysis of problem measurement situations.

RECOMMENDATIONS

It is your Committee's recommendation that the standards defined in this report become the objectives to which all practitioners should aspire, and which all users of program audience size measurements should demand.

It is your Committee's hope that the realization of these standards will be accelerated by the joint workings of technological progress in the field of program audience size measurement, the dynamic initiative of free research enterprise under the forces of competition, and by the enlightened leadership of the users of these measurements. Unless the Industry is prepared to pay the costs involved in such research, your Committee's standards will remain largely visionary.

The challenge before the Industry, particularly the practitioners, lies in improving and perfecting research technology to meet the standards set up in this report.

The degree of success that the various practitioners have in meeting these standards efficiently will be a measure of the value of what they have to offer to the Industry.

CHAPTER I •

Standards for Radio and Television Program Audience Size Measurements

This chapter presents your Committee's recommendations for standards in radio and television program audience size measurements. The Committee has proceeded on the theory that what is needed are measurements which will meet the greatest need of the greatest number of users.

These standards represent an optimum compromise, considering present theory, practice and cost. These standards are thought to be theoretically sound, and at the same time wholly realistic and practical.

Our standards are long-term objectives, not all of which can be met at the present time. They are presented as the goals toward which we hope program audience size measurement practitioners will aspire.

The program audience size measurement standards are grouped into three categories: information, procedure and accuracy.

The information standards are further sub-divided into basic and supplementary groups. The basic ones are those considered to be of primary importance. The supplementary ones are those which, while considered important, are extensions of the basic information standards. If there is a choice between them, the basic information standards should be given preference.

All procedure and accuracy standards are considered to be basic. By emphasizing the basic standards, the Industry will make substantial progress toward achieving uniformity in program audience size measurements. Your Committee considers this essential for the reduction of confusion in this field.

INFORMATION STANDARDS

Here your Committee states what information should be supplied by a program audience size measurement. The basic standards are discussed first; the supplementary ones next.

A · BASIC INFORMATION STANDARDS

1. Exposure to a broadcast should be measured in terms of set tuning

There are various levels of attentiveness which could serve for the definition of an "audience." On the one extreme there is the minimum requirement of set tuning. By set tuning we mean that a set be both turned on and receiving the program for which an estimated program audience size measurement is being obtained. On the other extreme one could require that the exposed individual be giving his undivided attention to the broadcast. This could be called an "attention" level of exposure. Between these two extremes there are various levels, such as:

a) "attended sets"—which could mean all tuned in sets that have one or more persons physically present.

b) "listening or viewing"—which could be the subjective opinion of the respondent as to whether or not he was paying attention to a program.

All of these are different levels of a condition which could be called "exposure."

Your Committee recommends that the Industry adopt the concept of tuning as its standard of exposure. This is the most objective of the various levels of exposure. It is also the simplest and most understandable level of exposure. It is the most inclusive of all exposure measures. It is the only measurement which does not require a subjective evaluation of some kind on the part of the exposed person (which evaluation is different for different respondents under identical exposure conditions). It is thus the only measurement which is unambiguous since it can only be interpreted one way.

In contrast, all other measurements are more subjective. For example, listening means different things to different people. Such ambiguity leaves much to the option of the respondent in reporting whether or not he is part of the "audience."

The possibility of defining audience size in terms of attended sets, as a compromise, was considered. However, your Committee could not define a satisfactory standard of what constitutes an attended set. For example, is a set attended to when someone is within ten feet of it in the same room but not attended to when someone is within ten feet of it in another room but still listening? Since attended sets cannot be defined

satisfactorily, tuning is recommended as the standard for the Industry for the basic program audience size measurement.

2. The unit of measurement should be the household

Program audience size measurements could be defined in terms of numbers of receivers, numbers of households, or numbers of individuals. The household rather than the individual is recommended as the basic standard for a program audience size measurement. In this definition a household is considered to be exposed to a program if at least one set associated with that household is exposed to the program.

A program audience size measurement reported on a household base has wider application than a measurement based upon individuals. This is because the household is the typical economic unit in marketing; and because the interaction between members of the household plays an important role in influencing most purchase decisions.

A measurement based on receivers is considered to be of lesser importance than households or individuals because receivers do not represent an economic unit.

3. All sets owned by the household should be measured

Your Committee recommends that the program audience size measurement should reflect the tuning of every household set regardless of its location. This concept includes both in- and out-of-home tuning, whether it is in the home itself, in an automobile, or a portable set used away from the household, just so long as the set is associated with the household being measured. This definition excludes such tuning as that which occurs in bars, restaurants and places of employment since these sets are not part of a household universe.

This definition of tuning as it relates to a household base is recommended for two reasons: (1) It restricts the definition of the total audience to a readily defined base; (2) The cost of measuring tuning of receivers not associated with households is usually not commensurate with the gain in program audience measurement coverage.

This definition does not give a program credit for coverage of additional households reached through individuals who are not members of the measured household. Measurement of non-household member expo-

sure is incompatible with the concept of program audience size measurements through household set tuning. It is also thought to be of nominal importance in the total audience activity.

4. The entire reception area should be measured

A program audience size measurement can cover the entire reception area of the station or stations carrying the program or it may cover some part of that total, such as the metropolitan area of the market or markets in which the broadcasting facilities are located. The former is preferred since it will give a program full credit for its entire audience.

However, for local measurements, some stations obtain coverages over such large geographical areas that in some cases it may be necessary to omit some of the "fringe" coverage. In such cases, the area to be measured should be so defined as to cover at least 90 per cent of the exposure hours of *each* station located in the market.

Separate areas for radio and television measurements may be necessary in a given market.

5. The measurement should be representative of all households

Your Committee recommends that program audience size measurements cover all types of homes regardless of economic class, telephone ownership, or any other characteristic. Otherwise, the measurement cannot be expected to reflect tuning activity in all households, because it would be derived from a biased sample of the area covered.

6. The measurement should report the average instantaneous audience

Program audience size measurements can estimate the average instantaneous audience for the duration of the entire program or the total audience tuned to some minimum part of the program.

Your Committee recommends the average instantaneous audience size measurement[1] as the standard because this measurement auto-

[1] This is a different concept from what might be called the average total audience measurement, which is obtained by first totaling exposure to each 15-minute segment and then dividing this total by the number of 15-minute segments in the program span.

matically weights persons or households into the audience in proportion to the amount of their tuning. For this reason, the average instantaneous audience measurement permits uniform comparisons of audience size measurements for programs of different duration.

A total audience measurement, while counting all households which were exposed over an arbitrary minimum of time, counts them all equally, regardless of how long they were exposed over that minimum. For this reason, a total audience size measurement will not permit comparison of audience size measurements for programs of different duration.

Further, acceptance of average instantaneous audience as the standard automatically makes the objective unambiguous. A total audience measurement requires an arbitrary decision as to what minimum amount of tuning to accept for including a household in the broadcast audience. Your Committee recognizes the value of a total audience measurement as the only measurement which reveals the full size of a program's audience, but designates it as a supplementary standard.

7. The measurement should express the number of households reached

Program audience size measurements can be expressed either as numbers or as percentages of some total. Your Committee recommends that all such measurements be reported in terms of number of households tuned in rather than in percentages of households.

If program measurements are expressed as percentages, a particular program can be given different estimates (even though these estimates represent the same absolute number of homes) because different bases may be used for the calculation of the percentages. Reporting in terms of numbers of households tuned in eliminates this source of variation. Audience size measurements for less than full network programs should also express the number of households reached.

Additionally, the number of homes reached is a fundamental measurement from which other important analyses can be derived; these include cost per thousand, percentages, etc.

B · SUPPLEMENTARY INFORMATION STANDARDS

While the basic program audience size measurement standards outlined in the preceding paragraphs would be the most important information provided, there is other valuable information which should be available. Such information does not need to be regularly reported.

8. Total household audience

This standard gives additional program audience size information by offering an estimate of the cumulative number of households which have tuned to any part, over some specified minimum, of a program broadcast.

9. Unduplicated household audience to two or more broadcasts

This information can provide cumulative total number of households reached over a given period of time, plus the average frequency with which the average household is reached. The information would be useful in determining the degree to which a program or group of programs reach the same or new household audiences.

The unduplicated household audience can be provided only on the basis of some type of total audience size measurement.

10. Program audience size measurements as per cent of all households

For analytical purposes such as trends and relative popularity, it is often useful to have program audience size measurements expressed in percentages. Your Committee, therefore, recommends as a supplementary measurement a program audience size measurement for each program as a per cent of all households.

For national program audience size measurements, this percentage would be based on all households in the United States.

For local program audience size measurements, it is necessary to define the area over which these measurements will be made, in order to obtain the base for the percentages.

Your Committee recommends, as one workable basis, the following procedure to define the limits of a market area in which to obtain program audience size measurements:

a) List all counties in which stations located in the market obtain exposure.

b) Arrange the list of counties for each station in descending order by estimated number of exposure hours.

c) Measure all counties necessary to obtain approximately 90 per cent of the exposure hours for *each* station.

d) Should the next county on the list account for approximately 5 per cent or more of the exposure hours of a station, your Committee recommends its inclusion also.

This process applied to all stations in the market will yield the counties to be included in the measurement area for that market. This one area should be used for per cent of household program audience size measurements for all stations in the market.

11. Households using receivers

For every measured 15-minute time period an estimate of the number of households in which at least one receiver[2] was in operation should be provided.

For local program audience size measurements, the area in which the number of households using receivers should be measured is that defined in the fourth and fifth paragraphs of Standard 10. In those instances where that area is identical with the basic program audience size measurement area (see Standard 4), a "share of audience" estimate can be computed in each report by dividing the number of homes reached by a given program by the number of households using receivers during the average minute that the program is on.

Where the area used for the basic program audience size measurement is not identical with that defined in Standard 10, shares of audience can only be computed when the information discussed in Standard 10 is made available, because only then are the program audience size measurement estimate and the estimate of households using receivers on the same base.

However, the share of audience thus computed for the less powerful stations would understate their competitive positions in the areas which

[2] In this section, "receivers" should be taken to mean "television sets" for television programs, "radio sets" for radio programs.

they serve, since the base for the computations would include areas they do not serve. In such instances, these less powerful stations might, on occasion at least, obtain special tabulations over lesser areas, provided the program audience size measurement sample is adequate to permit such computations.

Your Committee recommends that all national program audience size measurement reports include what is known as a "time period sets-in-use" (households using receivers) for each 15-minute period reported. This would be a simple count of the number of homes anywhere in the United States with receivers tuned in to any station during the average instant of the 15-minute period.

For network programs, share of audience computations would not be possible—since the program audience and the households using receivers would have been obtained for different periods in at least some of the markets—except in the instances where a given program is broadcast at a given New York time hour in all markets in which it is broadcast.

12. Audience characteristics

Since the sales potential for many commodities varies with such characteristics of the audience as sex, age, etc., your Committee recommends that measurements of audience size within certain population subgroups be made available. Some of these sub-group estimates would be obtained on an individual basis, some on a household basis. These data should be expressed as absolute numbers (households or individuals).

Individual measurements which are considered important are:

a) sex

b) age

c) product ownership or use (individually-owned items)

Household measurements which are considered important are:

a) family size

b) income class

c) geographic area

d) city or county size (regional or national measurements only)

e) product ownership or use (household-owned items)

Your Committee recognizes that no individual measurement can be defined strictly on a tuning basis. For this supplementary standard, therefore, your Committee recommends that an "attended sets" concept be applied.

13. Program audience size measurements for specific segments of a program, such as commercials

Program audience size measurements for program segments should be available.

14. Full network audience by specific time periods

This standard is included to permit measurement of the audience to all the stations of a particular network at a given time, whether or not the network has sold its full facilities for one particular program.

PROCEDURE STANDARDS

15. The program audience size measurement should be based upon a probability sample

Since it is impractical to measure audience size by taking a complete census of the population, the alternative is to select samples of the population. Your Committee specifies that such samples be probability samples. Only this kind of sample is projectible to the total population studied. Probability samples eliminate biased judgment in the selection of respondents, and only probability samples permit estimates of the precision of sample measurements.

16. The program audience size measurement should be based upon the audience during a single week

For one-a-week programs, your Committee recommends that the program audience size measurement be derived from a single broadcast of the program reported upon. For "strip" programs, your Committee recommends that the program audience size measurement be based upon the average of all of the broadcasts of that program during the week of the measurement.

17. The measurement should be reported for each commercially sponsored segment of the broadcast

Program audience size measurements should be available for all individually sponsored segments of programs. For sustaining programs the measurement should be available for the total program or customary segments. These will usually be for 15 minutes or multiples thereof.

18. All broadcast hours from 6 a.m. to midnight should be measured

This recommendation is based upon the fact that the overwhelming portion of broadcast time is sold between these hours. Measurements outside of this time span should be obtainable on special request.

19. The measurement should be available at varying frequencies, depending upon the importance of the market[3]

The high costs of network programs and competitive pressures make it important that program audience size measurements be available frequently. Only the collective needs of the Industry can determine this frequency. As one workable basis, your Committee suggests that national program audience size measurements and those covering major marketing areas be available monthly; those for secondary areas, quarterly; for smaller areas, it would be practicable to have one measurement covering a summer period and another one covering a period either during the fall, winter or spring.

This recommendation is based on a full consideration of the needs of sponsors, agencies and station-networks, of the research organization's physical and mechanical problems in processing data and of the costs of these measurements to the buyers. The recommended frequency is thought to be the optimum reconciliation of these three factors.

20. The measurement should be available within one month of the last measured broadcast[4]

Except for important changes in technology, demands for greater speed by users of program audience size data can only be met by in-

[3-4] These frequency and speed suggestions apply largely to the basic program audience size measurement standards. For most supplementary data, reports should be available less frequently as Industry demands indicate.

creases in costs and/or decreases in quality. It is to the Industry's self-interest to set up reasonable standards for speed of reporting to permit necessary time for the processing and quality controls without the burden of high costs due to overtime and related problems. Again, as one workable basis, your Committee suggests that all program audience size measurements be available within one month of the last measured broadcast.

ACCURACY STANDARDS

Your Committee recognizes that the total survey error in program audience size estimates results from the joint operation of two types of causes: (1) Sampling error—i.e., error arising from the use of a sample of the population rather than the entire population; (2) Non-sampling errors—i.e., all other errors—these can be classified under the following headings: Non-response; Conditioning of tuning behavior; Reporting or recording errors; Processing errors.

In the following two standards, your Committee specifies the amount of each type of error which it feels will be acceptable. These two standards should not be considered separately. Actually, it is your Committee's intention in setting these two standards to specify the total acceptable error in audience size surveys. In the two following standards we have specified one such total. Actually, your Committee would readily accept the results of any audience size measurement study whose *total* error was less than we have specified, no matter how the error was divided between sampling and non-sampling error.

21. There should be adequate control of sampling errors

Sampling error can be controlled by specifying the design and size of the sample. For the simplest type of probability sample—an unrestricted random sample—your Committee recommends a minimum sample size of 400 households for local audience size measurements and 1,200 for national audience size measurements for each 15-minute program or program segment. When other probability designs are used, the sample size should be adjusted to keep the sampling error within the limits specified above.

See Chapter IV for fuller exposition on sample size standards.

22. The net effect of non-sampling errors should not exceed the sampling error

(a) *Non-response.* In practice, no survey ever succeeds in obtaining usable replies from all designated households.

The principal reasons for failure to achieve a 100 per cent completion rate are:

(1) Some households refuse to cooperate, or are unable to do so due to language difficulties or other considerations.

(2) Some designated households or individuals simply cannot be reached, in spite of repeated call-backs (or because of absence of call-backs).

(3) Faulty sample implementation—due to interviewers' failure to follow instructions.

(4) In panel operations, drop-outs occur due to households moving, tiring of participation, and other reasons—thus reducing representativeness regardless of original placement rate.

(5) For telephone coincidentals, such things as busy signals and lines out of order prevent a 100 per cent completion at time of original call.

(b) *Conditioning of tuning behavior.* In some types of program audience size measurements, the sample members know that their behavior is being measured. This knowledge may cause them to change their behavior from what it would normally have been. In such cases, even if the respondents were chosen by a probability sample and their behavior correctly measured, this behavior would not be representative of the total universe being measured.

(c) *Reporting or recording errors.* Some of the errors which could contribute to these biases and which should be minimized are:

(1) Faulty keeping of diary records.

(2) Inability of one respondent to report accurately for every member of the family and for every receiver in the home.

(3) Mechanical failures.

(4) Memory failures.

(5) Prestige bias.

(6) Reporting or recording errors due to carelessness, lack of cooperation, lack of understanding and language difficulties.

(d) *Processing errors.* Good research practice stipulates that processing errors should be minimized by an adequate system of control, supervision and checking. These errors include those which can occur due to human and mechanical failure or errors during the processing of the raw research data into a final report. They include errors of editing, of computations, of formulae and of typing.

CHAPTER II •

The Program Audience Size Measurement Methods

This chapter is in two parts. The first part enumerates and describes seven fundamental methods now known for obtaining program audience size measurements. It also enumerates and describes three combination methods, each of which is based on two of the fundamental methods currently practiced. The second part names the more active commercial organizations which practice each of these methods on a continuing syndicated basis.

THE METHODS

The fundamental program audience size measurement methods are:

1. The Diary
2. The Recorder
3. The Personal Coincidental
4. The Personal Roster Recall
5. The Personal Unaided Recall
6. The Telephone Coincidental
7. The Telephone Recall

The combination methods[1] are:

8. The Combination Telephone Coincidental and Telephone Recall
9. The Combination Telephone Coincidental and Diary
10. The Combination Telephone Coincidental and Personal Roster Recall

[1] The A. C. Nielsen Company recently announced its plans for syndicating a local program audience size measurement service which is a combination of the recorder and diary methods. Because this method was not being practiced commercially, at the time this report was written, it will not be discussed.

A brief description of each of these methods follows:

The Diary. This method requires that some member (or members) of the household keep a written record or log of program exposure.

The Recorder. This is a method which electronically or mechanically records automatically individual set tuning including frequency or channel to which the set is tuned.

The Personal Coincidental. Personal interviews are made throughout the duration of a given program or time period, and respondents are queried regarding program exposure in the household at the moment of call.

The Personal Roster Recall. Respondents are shown a list of programs and stations and asked to indicate which they were exposed to during the measured time span.

The Personal Unaided Recall. Personal interviews are made during which respondents are asked about program exposure for a preceding span of time. Unlike the roster, the personal unaided recall uses no list of programs or stations, depending entirely upon the respondent's unaided memory for exposure information.

The Telephone Coincidental. This method employs the same principles as the personal coincidental method except that the interviews are made by telephone.

The Telephone Recall. This method employs the same principles as the personal unaided recall except that the interviews are made by telephone.

The Combination Telephone Coincidental and Telephone Recall. Respondents are asked in a telephone interview about both coincidental and previous broadcast exposure.

The Combination Telephone Coincidental and Diary. This method combines broadcast exposure information obtained by the coincidental telephone method in one sample of homes with information obtained by the diary method in another sample of homes.

The Combination Telephone Coincidental and Personal Roster Recall. This method combines broadcast exposure information obtained by the telephone coincidental method in one sample of homes with information obtained by the roster recall method in another sample of homes.

THE PRACTITIONERS

The following table lists the more active organizations which practice the methods discussed in this chapter.

Methods	Practitioners
The Diary	American Research Bureau, Inc. Videodex, Inc.
The Recorder[2]	A. C. Nielsen Company
The Personal Coincidental	(No continuing service)
The Personal Roster Recall	The Pulse, Inc.
The Personal Unaided Recall	(No continuing service)
The Telephone Coincidental	Trendex, Inc.
The Telephone Recall	(No continuing service)
The Combination Telephone Coincidental and Telephone Recall	C. E. Hooper, Inc.
The Combination Telephone Coincidental and Diary	C. E. Hooper, Inc.
The Combination Telephone Coincidental and Personal Roster Recall	Robert S. Conlan & Associates

For further information on the operations of each of the practitioners, the reader is referred to ARF's *Directory of Audience Size Measurement Services,* published in January, 1954.

[2] C. E. Hooper, Inc., and The Pulse, Inc., have notified your Committee that a recorder program audience size measurement service will be placed in commercial practice. Because these services were not being practiced commercially, at the time this report was written, C. E. Hooper, Inc., and The Pulse, Inc., are not shown as practitioners.

CHAPTER III ·

The Potential of Program Audience Size Measurement Methods

This chapter outlines your Committee's judgment as to how well each of the radio and television program audience size measurement methods would meet our standards if practiced to its fullest potential. By fullest potential we mean the optimum compromise among theory, practice and cost. This optimum compromise accepts neither the "ivory tower" ideal on the one hand nor cheap research which uses spurious short cut methods on the other. It is based upon methodological principles which are accepted at the highest levels of research practice today.

It is again reiterated that the discussion which follows is concerned only with the *methods* themselves when practiced to their fullest potential. No statement is made, or should be inferred, as to how closely present practice comes to realization on any given method's fullest potential. Such an analysis is planned by your Committee.

The program audience size measurement methods are evaluated for each standard independently, without any consideration of the possible interrelationships among the standards themselves. If this were not done, failure of a method to meet any one standard might preclude its meeting any others. For example, because a method cannot deliver an average instantaneous estimate, it does not mean it cannot deliver a representative sample of all households or an estimate of all set tuning. Your Committee's procedure of analysis permits evaluation of each method on every standard which has been established, thus giving a complete evaluative profile for each method.

In the table, a qualified response is sometimes stated for a given method. For example, under the heading "Reporting or Recording Errors" for the recorder, the phrase "yes, if mechanical failures negligible" is used. In this instance, your Committee thinks there can be adequate

control of all other non-sampling errors based upon the present level of technology.

In other instances, a "no" could, theoretically, have been stated as a "qualified yes." In such instances, however, we are saying "no" because your Committee thinks that the standard is unattainable in practice at the present level of technology.

The table at the end of this chapter will summarize the potential of each method when practiced to its fullest potential. The details are discussed in the following text.

INFORMATION STANDARDS

A · BASIC INFORMATION STANDARDS

1. Exposure to a broadcast should be measured in terms of set tuning

All methods can produce estimates of set tuning.

2. The unit of measurement should be the household

All methods can report their program audience size measurements upon a household audience base. However, as stated in the next section, there may be significant differences in the extent to which each method can report upon all sets within the household measured.

3. All sets owned by the household should be measured

All methods can meet this standard for television since there are no household sets outside the home. All methods cannot meet this standard for radio. The diary, the roster and the unaided recall methods (telephone and personal) can deliver estimates of total radio exposure related to a household base. All other methods fall short of this objective. The recorder cannot produce estimates of exposure to all radio sets in the household because it cannot measure exposure to battery and portable sets.

The personal coincidental methods similarly cannot deliver an estimate of total exposure on a household base for radio because of the practical impossibility of coincidental measurement of out-of-home activity. Two of the combination methods—the combination telephone coinci-

dental and diary and the combination telephone coincidental and personal roster recall—can produce an estimate of audience size covering all household sets because the diary can extend the telephone coverage to all sets in the first combination and the personal roster recall can extend the telephone coverage in the second combination. No such extension is possible for the combination telephone coincidental and telephone recall.

4. The entire reception area should be measured

All methods can deliver an audience size estimate based upon the total area covered by the station(s) carrying the program. However, there can be large differences in the overall costs of obtaining this coverage.

In general, the lowest unit costs in obtaining area coverage and household representativeness are associated with those methods which get the largest amount of information per unit of contact cost. On one extreme would be the personal coincidental method where the one contact with a household gives only one unit of exposure information—that exposure occurring at the moment of the call. On the other extreme would be panel operations which, while requiring a larger initial investment per original contact, can amortize the investment over the information obtained for literally thousands of broadcasts.

Because national program audience size measurements offer practitioners a substantially greater revenue potential than any individual local program audience size measurement and because the increased costs of obtaining national program audience size measurements grow at a slower rate than potential revenues, the economics of local methods will sometimes discourage the introduction of methods requiring large initial investments for obtaining local program audience size estimates.

5. The measurement should be representative of all households

All methods can produce a sample which represents all households in the universe surveyed. However, in the case of methods based upon telephone homes only, a sample may be unrepresentative if telephone ownership does not approach saturation. The combination telephone coin-

cidental and diary can do this through the diary. The combination telephone coincidental and personal roster recall can do this through the roster.

6. The measurement should report the average instantaneous audience

Of the fundamental methods, the recorder, the personal coincidental and the telephone coincidental can deliver an average instantaneous audience estimate. For all practical purposes, all other fundamental methods can only deliver a total audience measurement, although theoretically, an open-end diary[1] could produce an approximation to an average instantaneous audience. The combination methods can do it through adjustment by the coincidental part of the measurement.

7. The measurement should express the number of households reached

All methods can produce program audience size estimates expressed in numbers of households reached for the universe surveyed.

B · SUPPLEMENTARY INFORMATION STANDARDS

8. Total household audience

Of the fundamental methods, the diary, the recorder and the recall methods can deliver estimates of total audience. The personal and telephone coincidental methods cannot do this because they cannot cumulate audience. All of the combination methods can produce a total household audience through adjustment.

9. Unduplicated household audience to two or more broadcasts

Only the diary, the recorder and the combination telephone coincidental and diary can deliver estimates of unduplicated audiences over a number of programs and a number of broadcasts of the same program. However, the degree to which these methods can do this is limited

[1] In open-end diaries, the respondent could be asked to record to the nearest minute tune-in and tune-out for each program. However, this cannot, in practice, be carried out with sufficient accuracy.

by the length of time sufficient members of the fixed panel report without interruption.

For practical purposes, none of the other methods can produce unduplicated audience measurements for any long period of time.

10. Program audience size measurements as per cent of households

All methods can produce program audience size measurements expressed as percentages of all households for the universe surveyed.

11. Households using receivers

All methods can produce some type of a households using receivers measurement.

12. Audience characteristics

All methods can report program audience size measurements based upon household characteristics, such as family size, income class, product ownership or use and geographic location.

Program audience size measurements based upon characteristics of the individuals exposed to a program, such as age and sex, can be provided by all methods except the recorder which is restricted to measuring whether the set is turned on or off and the station to which tuned.

13. Program audience size measurements for specific segments of a program, such as commercials

Only the recorder can deliver reliable estimates of audience size to program segments, such as commercials.

While the personal and telephone coincidentals can theoretically provide such estimates, they can only do it commercially at prohibitive costs because of sampling requirements. Specifically in a telephone or personal coincidental, 400 local or 1,200 national interviews would have to be made during each segment to meet your Committee's sample size requirements. If we assume that it would be desirable to measure segments (commercials, for example) one minute long, this would mean

a minimum sample of 6,000 locally and 18,000 nationally for these services for a 15-minute program.

While non-coincidental methods can attempt to make estimates of audience size to program segments, your Committee thinks that the quality of such estimates would be substantially inferior to their estimates of program audience size as a whole. Since the quality of such measurements is inferior, your Committee rejects them as estimates of audience size.

14. Full network audience by specific time periods

All methods can deliver a full network program audience size measurement by specific time periods.

PROCEDURE STANDARDS

15. The program audience size measurement should be based upon a probability sample

Probability sample designs can be employed with all methods. However, probability samples can only be applied to the universes measured by the method. For example, the telephone methods can apply probability designs to telephone home samples; their failure to cover all homes is discussed in Standard 5, "The measurement should be representative of all households."

However, a probability sample design applied to a fixed panel tends to become biased through time because of drop-outs and changes in the composition of the universe covered. A fixed panel operation must therefore develop some technique for maintaining probability design through time.

16. The program audience size measurement should be based upon the audience during a single week

All methods can produce program audience size measurements based on a single broadcast of the program measured (or the average of all broadcasts in the case of "strip" programs).

17. The measurement should be reported for each commercially sponsored segment of the broadcast

All methods can produce program audience size estimates for commercially sponsored segments.

18. All broadcast hours from 6 a.m. to midnight should be measured

Because persons cannot be reached at all hours of the day, either by telephone or by personal interview, the coincidental methods and combination methods wholly dependent on coincidental methods cannot measure exposure over all broadcast hours between 6 a.m. and midnight. All other methods can report on all broadcast hours.

19. The measurement should be available at varying frequencies, depending upon the importance of the market

Theoretically, all methods can deliver program audience size measurements at any desired frequency provided the Industry is willing to underwrite the cost.

20. The measurement should be available within one month of the last measured broadcast

All methods can deliver program audience size measurements with a relatively high degree of speed. Theoretically, the fastest reporting could be based on an electronic principle which reduces the time required to collect exposure data by centrally recording it as it occurs. The coincidental and recall methods are next in point of speed, followed by the diary.

ACCURACY STANDARDS

21. There should be adequate control of sampling errors

All methods can deliver program audience size estimates based upon adequate sample size and design, both at the local and national levels. However, some methods can do this more economically than others. This becomes particularly important at the local level when the total amount of money the Industry is willing to invest in program audience

size measurement research is incommensurate with the high cost of sampling. This limitation is particularly applicable to recorder methods.

22. The net effect of non-sampling errors should not exceed the sampling error

a) *Non-response.* All methods except the personal coincidental can limit non-response sufficiently to meet this standard.

The personal coincidental cannot meet this standard because of the high refusal rate at the door and because call-backs are incompatible with the method.

The telephone coincidental meets this standard because of the higher completion rate on the telephone than at the door. Other methods can overcome at least a part of refusals at the door through call-backs.

b) *Conditioning of tuning behavior.*[2] Ideally, the program audience size measurement method itself should have no effect upon tuning behavior. That is, the method itself should not effect either program choice or the duration of exposure.

Of the methods described, only the diary tends to exercise a conditioning effect upon exposure. None of the other methods except the recorder could exercise any conditioning effect upon tuning behavior, because tuning precedes the awareness of having to report. The recorder is thought to exercise negligible effect on exposure behavior because it can be almost completely unobtrusive.

c) *Reporting or recording errors.* Only the recorder minimizes errors of reporting or recording, provided mechanical failures are negligible, because of its freedom from human error.

The diary is subject to inaccurate and incomplete recording on the part of the persons keeping it. Prestige bias and memory errors are of increasing importance as the interval between exposure and recording increases. This is due to the fact that the reliability of respondent testimony decreases as the interval between exposure and the reporting of

[2] Your Committee distinguishes between errors of reporting or recording and those due to conditioning of tuning behavior. Errors of reporting or recording have to do with human or mechanical failures to report what actually occurred. Conditioning of tuning behavior reflects the possible effect of the measuring technique itself upon people's actual tuning behavior.

that exposure increases. Memory failure can have either an upward or downward bias due to forgetting and confusion.

The recall methods are subject to prestige bias, incomplete reporting and memory failure, accentuated here by unequal intervals between exposure and recording at different times of the day.

Errors of reporting are associated with all personal and telephone methods since the person or persons reporting for the household may not be aware of the exposure of other family members not in the immediate presence of those reporting for that household.

d) *Processing errors.* All methods can reduce processing errors to a minimum through adequate supervision and checking.

CONCLUSION

The foregoing analysis indicates that there are strengths and weaknesses in all the methods. Each method meets some but not all of the standards.

The challenge before the Industry, particularly the practitioners, lies in improving and perfecting research technology to meet the standards set up in this report.

The degree of success that the various practitioners have in meeting these standards efficiently will be a measure of the value of what they have to offer to the Industry.

THE POTENTIAL OF RADIO-TV PROGRAM AUDIENCE SIZE MEASUREMENT METHODS UNDER HIGHEST STANDARDS OF RESEARCH IMPLEMENTATION

(Based on present known technology)

	DIARY	RECORDER	PERSONAL COINCIDENTAL	PERSONAL ROSTER RECALL	PERSONAL UNAIDED RECALL	TELEPHONE COINCIDENTAL	TELEPHONE RECALL	COMBINATION TELEPHONE COINCIDENTAL AND TELEPHONE RECALL	COMBINATION TELEPHONE COINCIDENTAL AND DIARY	COMBINATION TELEPHONE COINCIDENTAL AND PERSONAL ROSTER RECALL
INFORMATION STANDARDS*										
A. Basic Information Standards										
1. Exposure to a Broadcast Should Be Measured in Terms of Set Tuning	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2. The Unit of Measurement Should Be the Household	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3. All Sets Owned by the Household Should Be Measured	Yes	Yes for TV No for radio—does not measure battery and portable sets	Yes for TV No for radio—measurement of out-of-home exposure practically impossible	Yes	Yes	Yes for TV No for radio—measurement of out-of-home exposure practically impossible	Yes	Yes for TV No for radio—measurement of out-of-home exposure practically impossible	Yes	Yes
4. The Entire Reception Area Should Be Measured	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5. The Measurement Should Be Representative of All Households	Yes	Yes	Yes	Yes	Yes	Yes—if telephone ownership approaches saturation	Yes—if telephone ownership approaches saturation	Yes—if telephone ownership approaches saturation	Yes	Yes
6. The Measurement Should Report the Average Instantaneous Audience	No—in a practical sense only a total program audience size measurement is possible	Yes	Yes	No—in a practical sense only a total program audience size measurment is possible	No—in a practical sense only a total program audience size measurement is possible	Yes	No—in a practical sense only a total program audience size measurement is possible	Yes—if adjusted	Yes—if adjusted	Yes—if adjusted
7. The Measurement Should Express the Number of Households Reached	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
B. Supplementary Information Standards										
8. Total Household Audience	Yes	Yes	No—measures only average instantaneous audience	Yes	Yes	No—measures only average instantaneous audience	Yes	Yes—if adjusted	Yes—if adjusted	Yes—if adjusted
9. Unduplicated Household Audience to Two or More Broadcasts	Yes	Yes	No—measures only average instantaneous audience	No—except for periods of roster interview	No—except for periods of roster interview	No—measures only average instantaneous audience	No—except for periods of roster interview	No	Yes	No
10. Program Audience Size Measurements as Per Cent of All Households	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
11. Households Using Receivers	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12. Audience Characteristics a) Individual b) Household	Yes Yes	No Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes

* The audience size measurement methods are evaluated for each standard independently without any consideration of the possible interrelationships among the standards themselves.

THE POTENTIAL OF RADIO-TV PROGRAM AUDIENCE SIZE MEASUREMENT METHODS UNDER HIGHEST STANDARDS OF RESEARCH IMPLEMENTATION (Continued)

(Based on present known technology)

	DIARY	RECORDER	PERSONAL COINCIDENTAL	PERSONAL ROSTER RECALL	PERSONAL UNAIDED RECALL	TELEPHONE COINCIDENTAL	TELEPHONE RECALL	COMBINATION TELEPHONE COINCIDENTAL AND TELEPHONE RECALL	COMBINATION TELEPHONE COINCIDENTAL AND DIARY	COMBINATION TELEPHONE COINCIDENTAL AND PERSONAL ROSTER RECALL
13. Program Audience Size Measurements for Specific Segments of a Program, Such as Commercials	No	Yes	No—sample requirements prohibitive	No	No	No—sample requirements prohibitive	No	No	No	No
14. Full Network Audience by Specific Time Periods	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PROCEDURE STANDARDS*										
15. The Program Audience Size Measurement Should Be Based Upon a Probability Sample	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
16. The Program Audience Size Measurement Should Be Based Upon the Audience During a Single Week	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
17. The Measurement Should Be Reported for Each Commercially Sponsored Segment of the Broadcast	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
18. All Broadcast Hours from 6 A.M. to Midnight Should Be Measured	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes
19. The Measurement Should Be Available at Varying Frequencies, Depending Upon the Importance of the Market	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
20. The Measurement Should Be Available Within One Month of the Last Measured Broadcast	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ACCURACY STANDARDS*										
21. There Should Be Adequate Control of Sampling Errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
22. The Net Effect of Non-Sampling Errors Should Not Exceed the Sampling Error										
a.) Non-Response	Yes	Yes	No—the relatively low completion rate cannot be reduced by call-backs	Yes—although all exposed not available on first and subsequent calls	Yes—although all exposed not available on first and subsequent calls	Yes—with completion rate higher than personal methods	Yes—although all exposed not available on first and subsequent calls	Yes—with completion rate higher than personal methods	Yes	Yes
b.) Conditioning of Tuning Behavior	No—may affect both type and amount of exposure	Yes—minimizes exposure behavior	Yes—interview cannot affect past exposure behavior	Yes—interview cannot affect past exposure behavior	Yes—interview cannot affect past exposure behavior	Yes—interview cannot affect past exposure behavior	Yes—interview cannot affect past exposure behavior	Yes—interview cannot affect past exposure behavior	No—contains diary biases	Yes—interview cannot affect past exposure behavior
c.) Reporting or Recording Errors	No—subject to inaccurate and incomplete recording	Yes—if mechanical failures negligible	No—subject to respondent reporting problems	No—subject to memory failure and prestige bias and incomplete reporting	No—memory and prestige biases maximized; subject to incomplete reporting	No—problems of incomplete reporting for all household members	No—subject to memory failure, prestige bias and incomplete reporting	No—problems of incomplete reporting	No—has telephone and diary biases	No—has errors of telephone coincidental and personal recall
d.) Processing Errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* The audience size measurement methods are evaluated for each standard independently without any consideration of the possible interrelationships among the standards themselves.

CHAPTER IV ·

Statistical Considerations

From the standpoint of time and cost, it is impossible to take a complete census of every household in a universe to obtain program audience size measurements. In practice, therefore, these measurements are derived from samples of the population in the universe. This raises such questions as how large the samples should be and how they should be drawn.

In earlier chapters your Committee presented its recommendations on sampling procedures, stating its preference for probability samples and specifying minimum sample size requirements. In the present chapter, these recommendations are discussed in more detail.

Your Committee recommended that only those samples be used for which the probability of inclusion of each household is known—i.e., probability samples. This recommendation is made because: (1) Program audience size measurement estimates based on such samples can be shown to be free of sampling bias if the sampling procedure is properly carried out; (2) Estimates made from such samples permit measurement of their reliability. Program audience size measurement estimates derived from non-probability samples meet neither of these tests. Consequently, any confidence in the reliability of program audience size measurement estimates based on non-probability samples must be based on faith alone.

SAMPLING ERROR

To clarify its sample size standards, your Committee discusses briefly three factors that affect the reliability of program audience size measurements:

1. *The type of probability sample.* There are many methods of selecting a true probability sample. Some of these methods are more efficient

than others in minimizing sampling error. For example, the simplest design—unrestricted random sampling—generally will be more efficient, in practice, than a more complicated design which involves clustering (even though the actual number of "interviews" may be the same in the two samples). Your Committee cannot possibly comment on all types of probability samples which might be used in program audience size measurements. We have specified therefore that the samples used should provide reliability "equal to that provided by samples of 400 households locally and 1,200 households nationally selected by unrestricted random sampling." Any practitioner using a probability sample will be able to tell whether his sample meets this standard.

2. *Size of the sample.* As sample size is increased, sampling error decreases. However, to cut sampling error in half, sample size must be multiplied by four. Similarly, to cut sampling error to a third, sample size must be multiplied by nine. Such increases in sample size would be expensive. Your Committee has tried therefore to avoid both extremes—too small samples with too little reliability, and too large samples with too high costs. It should be remembered that, no matter how large a sample your Committee had specified, there would still be some program audience size measurements which would be unsatisfactory because the relative sampling error was too high. This problem will generally arise in the case of broadcasts whose audience size is very small compared to the total universe. The problem can be partly solved by combining several weeks' measurements to get an overall program audience size measurement subject to less sampling error.

3. *Use of additional information.* Frequently, a sample design or estimate can be improved by the use of reliable outside information. Suppose, for example, that a television audience study is to be made in two counties—A and B. Suppose, further, that the following facts are established:

Homes in A	—100,000
Homes in B	—100,000
TV Homes in A—	50,000
TV Homes in B—	0

In this case your Committee's sample requirements would be met if, without resort to the available outside information, 400 sample homes were spread throughout A and B by unrestricted random sampling. How-

ever, the sample estimate would be subject to much less sampling error if we made use of the available outside information by placing all of the sample homes in county A.

The intelligent use of reliable outside information can improve sample estimates. However, the use of biased or unreliable outside information estimates in conjunction with good probability sampling methods will only serve to bias the results obtained or make them less reliable. In such cases, it is better to depend on probability sampling methods only and to avoid the use of the outside information.

SAMPLING RECOMMENDATIONS

Your Committee recommends the following sample sizes for program audience size measurements for two reasons: The diminishing returns involved in increasing the precision of program audience size measurements by increasing sample size; and the level of reliability obtained from these sample sizes will be adequate for most purposes.

For unrestricted random samples your Committee recommends a minimum sample size of: 400 households in final tabulation for each local program audience size measurement; and 1,200 households in final tabulation for each national program audience size measurement. Where several broadcasts are averaged (as in strip programs) and the same households are "interviewed" several times, the statistical equivalent of these requirements should be provided.

For probability samples in which households are not selected by unrestricted random sampling, your Committee recommends that the samples used provide reliability at least as great as would be provided by unrestricted random samples of 400 locally and 1,200 nationally.

Your Committee specifies that the minimum samples should be 400 (or 1,200) households. Even for television program audience size measurements, it is not necessary that all the cases be television homes because the sample should be projectible to the total households in the area being measured.

Your Committee has called for a smaller sample in local measurements than in national measurements. To get the same reliability in local program audience size measurements as in national measurements, the same sample size (1,200) would be required. However, less precision is re-

quired for local measurements. From a purely business point of view, the larger the investment in time and talent, the more accurate should the information be on which decisions are based.

It is recommended that reports of program audience size measurements be accompanied by estimates of their precision if probability sampling methods were used in obtaining the estimates. In the interest of uniformity all estimates of sampling error should be reported in terms of one standard error of estimate. This would not be necessary for every measurement in the report. It would be sufficient, for example, if a local report were to contain a table something like the following example:

TOTAL HOUSEHOLDS IN AREA MEASURED—1,000,000

Program Audience Size Estimates (no. of households)	Corresponding Ratings (per cent)	Standard Error of Estimate (no. of households)	ARF Maximum Sampling Error Standard (no. of households)
5,000	.5	2,822	3,527
10,000	1.0	3,980	4,975
20,000	2.0	5,600	7,000
30,000	3.0	6,823	8,529
40,000	4.0	7,838	9,798
50,000	5.0	8,703	10,897
60,000	6.0	9,499	11,874
70,000	7.0	10,206	12,757
80,000	8.0	10,852	13,565
90,000	9.0	11,447	14,309
100,000	10.0	12,000	15,000
150,000	15.0	14,283	17,854
200,000	20.0	16,000	20,000
250,000	25.0	17,321	21,651
300,000	30.0	18,330	22,913
350,000	35.0	19,078	23,848
400,000	40.0	19,596	24,495
450,000	45.0	19,900	24,875
500,000	50.0	20,000	25,000
550,000	55.0	19,900	24,875
600,000	60.0	19,596	24,495
650,000	65.0	19,078	23,848
700,000	70.0	18,330	22,913
750,000	75.0	17,321	21,651

Column 1 lists selected program audience size measurement estimates.

Column 2 lists the corresponding rating—or per cent of the assumed 1,000,000 households in the area measured.

Column 3 gives the standard error of estimate associated with the indicated program audience size estimates. This is a hypothetical example for the sample size and design employed by a program audience size measurement practitioner.

Column 4 gives the ARF maximum allowable sampling error—the sampling error associated with program audience size estimates based on an unrestricted random sample of 400 households.

By examining the third column of this table the user could determine the reliability of any particular program audience size measurement. By comparing the last two columns, he could determine whether the sampling reliability of the report met the ARF sampling error standard. Only if the figures in the third column are smaller than those in the fourth column is the sampling error standard being met.

TECHNICAL NOTES

1. The discussion in this and previous chapters has referred to unrestricted random sampling. Other types of probability sampling designs may have equal or even greater reliability than unrestricted random designs of the same sample size. However, the costs of implementing such ultra-efficient designs are usually prohibitive unless they can be amortized through time. This usually involves interviewing the same households again and again.

Most probability sampling designs involve clustering. This usually results in a sample design which is appreciably less efficient than an unrestricted random sample of the same size. In such cases, the minimum ARF standards (400 local, 1,200 national by unrestricted random sampling) may require sample sizes of perhaps two or three times those required for unrestricted random samples to produce equivalent reliability.

2. One of the consequences of using relatively small samples is that relatively large variations in program audience size measurements can be explained on the basis of sampling error and may not be necessarily due to actual variations in the universe being measured. When an important decision is to be made, it may be advisable for the buyer to base his decision on the average of several such program audience size measurements. This combination would be subject to a smaller sampling error than any of the individual measurements.

3. This chapter dealt with sampling error. There are other important factors which affect the accuracy of program audience size measurements. These non-sampling errors were discussed in Chapter I. Frequently it is more desirable to use additional funds to improve control of non-sampling errors, instead of using the money to obtain samples designed to produce statistically more reliable estimates.

CHAPTER V ·

Standards for Evaluation

This chapter is not presented as a criticism of present reporting practices in the field. It is presented in a constructive spirit because it is thought that the Industry must present uniform evaluation standards as another goal toward which services should aspire. Application of these standards for evaluation will make it possible to understand more fully the differences among program audience size measurements produced on the same program, at the same time, by various organizations.

Your Committee thinks it should be the responsibility of the Industry, rather than of each practitioner, to decide what information should be supplied for evaluation of a service. Such uniformity in reporting would tend to overcome confusion.

Your Committee recommends that each program audience size measurement practitioner should report at least a minimum amount of information to permit buyers to evaluate the technical bases upon which their measurements were obtained. This information should be of two broad types: (1) Data which report the scope and method of the study, outlined in detail in each report; (2) Reserve data available in the offices of the service, not necessarily published in each report. Each of these two bases is discussed next.

1. *Report on scope and method.* Where they apply, each practitioner can be guided by the *Criteria For Marketing and Advertising Research,* published by the ARF in April, 1953. The principles presented in the *Criteria* should be used as guides in outlining to the reader the nature of the study, its scope, design and implementation. Your Committee cannot stress too strongly the need for this thorough-going reporting on the methods and controls used. It is only by this form of full reporting of the technical methods involved that the strengths and weaknesses of each program audience size measurement method can be effectively gauged.

Without this information, buying of services frequently has to be made on bases which are not relevant to the subject matter.

The ARF *Criteria* listed the following questions which should be asked by the users of research in evaluating research studies:

1. Under what conditions was the study made?
2. Has the questionnaire been well designed?
3. Has the interviewing been adequately and reliably done?
4. Has the best sampling plan been followed?
5. Has the sampling plan been fully executed?
6. Is the sample large enough?
7. Was there systematic control of editing, coding and tabulating?
8. Is the interpretation forthright and logical?

Your Committee recommends that the program audience size measurement practitioners give adequate information in each report to permit the users of their reports to answer each of these questions satisfactorily (wherever answers are relevant). In particular, it is pointed out that sample size adequacy can only be determined with regard to the effective sample—that is, the number of households on which information is obtained. Accordingly, each report should give information on the effective sample size.

It is further recommended that the sampling reliability of the program audience size measurements in the report be expressed by means of a table similar to that suggested in Chapter IV. Where probability sampling methods are not used, such a table would be inapplicable. In such cases, the report should indicate that probability sampling methods have not been used.

2. *Reserve data.* It is frequently necessary to refer to the original data to clarify additional questions on program audience size measurements. Unless these original data are kept for a reasonable period of time, little can be done to clarify some questions.

It is recommended that each practitioner keep certain basic data from each study available for reference for at least one year after the program audience size measurements are published. These basic data (where applicable) include:

1. The number of homes with which contacts were attempted for any given time period for which data are reported.
2. The number of homes from which the information was obtained.
3. The names and addresses of respondents contacted, written on the original interviewing sheets.
4. The name of the interviewer, written on the original interviewing sheets.
5. Date of the interview.
6. All tabulation data.
7. The addresses of all interviewers who were employed for the survey.
8. Details of the sample plan to be used for reference purposes, stating how the sample was drawn and executed.
9. Instructions to interviewers for the period covered.
10. Supervisors' records of checks on interviewers' work.

Appendix A

GLOSSARY

The following terms which appear in this report are defined below in alphabetical order. In the definitions any word which has itself been defined is set in italics.

Accuracy: A measure of how closely a true *population* measurement is approximated by the survey estimate of the measurement.

Attended Set: A television or radio receiver that is both tuned to a *broadcast* and has at least one responsible individual present in a specified neighborhood of the set.

Audience (household basis): The number of *households* exposed to a *broadcast*. A household is said to be exposed to a broadcast if at least one set belonging to that household is exposed to the broadcast.

Audience (individual basis): The number of responsible individuals present in a specified neighborhood of any set exposed to a *broadcast*.

Average Instantaneous Audience: The average number of *households* (individuals) exposed to a program over the duration of the *broadcast*.

Bias: An error that results from an incorrect or inadequate procedure in design and implementation of a survey. The size of a bias is the numerical difference between the true *population* measurement and the expected value of the survey estimate of the measurement.

Broadcast: In this report broadcast denotes either radio or television transmission, or both, whichever is appropriate.

Call-back: An attempt by an interviewer to obtain cooperation from a designated respondent in a survey who was not interviewed on a previous attempt.

Conditioning of Response: A survey error arising from a respondent altering his *exposure* behavior as a result of finding out that his behavior is being, or will be, measured.

Coverage: A station's coverage is its potential *audience*. Coverage can be expressed in many ways. Three of the most common are:

1. Engineering contours. It is assumed that all radio or television equipped homes within range of an arbitrary minimum signal strength are "covered" by a station.
2. Minimum *listening*. A home, no matter where located, is assumed to be "covered" by a station if it has an arbitrary minimum amount of *exposure* to that station.
3. Minimum per cent of *exposure*. A county is said to be "covered" by a station if a minimum per cent of *households* report some minimum per cent of *exposure*.

Error of Reporting or Recording: Failure by either the respondent, interviewer, or recording instrument to report without error the characteristic being measured.

Exposure: A degree of attention to a *broadcast*, such as *listening*, set *tuning*, or presence at the set.

Exposure Hours: A station's exposure hours is the sum of all the time exposed to that station by all homes regardless of where they may be, for a given span of time.

Full Network Audience: The *audience* of all network stations during a given time period regardless of what program is carried on any one station during the given time period.

Household: A household includes all of the persons who occupy a dwelling unit. It includes the related family members and also the unrelated persons, if any, such as lodgers, maids or hired hands who share the dwelling unit. A person living alone or a group of unrelated persons sharing the same living accommodations as partners is counted as a household.

Household Set: Any radio or television receiver owned by members of a *household*. Such receivers include portable and automobile sets.

Households Using Receivers: The number of *households*, in a given area, with one or more receivers in use during a measured period. This could be defined separately for radio and television. It could also be defined in terms of the *average instantaneous audience* or *total audience*.

Listening or Viewing: A subjective opinion by a respondent who is in hearing range of a receiver as to whether or not he was paying attention to the program.

Market Coverage Area: All counties in the *station coverage area* of any station emanating from a given market.

Population or Universe: The totality of all elements of a certain kind, such as all individuals 10 years of age or over residing in the United States or all United States *households*.

Prestige Bias: A survey error arising from a respondent's attempt to upgrade the level of his cultural interests by incorrectly reporting the programs to which he is, or was, exposed.

Probability Sample: A *sample* chosen in such a way that the probability of the selection of any *population* element is known.

Rating: The per cent of all *households* (or individuals) in a *population* that is in the *audience* of a given *broadcast*.

Reception Area: The area over which a given station obtains its *coverage*.

Reliability: An evaluation of the influence of *sampling error* on a given *sample* measurement. The less the sampling error, the greater the reliability of the estimate.

Sample: A group of elements chosen from a *population* to represent the population.

Sampling Error: The error in an estimate that results from its being based on a *sample* instead of the *population* from which the sample was drawn. The size of the sampling error is the numerical difference between the sample result and the result that would have been obtained if the sample survey procedures were applied to every member in the defined *universe*.

Share of Audience: The per cent of *households using receivers* (either radio or television) which are exposed to a given program.

Standard Error of Estimate: A statistic used to measure the *reliability* of a *sample* estimate. A precise definition of this term is beyond the scope of this report; however, it is defined in all basic textbooks on statistics.

Station Coverage Area: The land area containing all or part of a station's *audience*. One such area could be defined to be those counties, minimum in number, in which 90 per cent of a station's total weekly *exposure hours* occur.

Strip Programs: Programs *broadcast* on the same station or group of stations at the same broadcast hour two or more days of the week.

Total Audience: The number of *households* (or individuals) that are exposed to any part of a program over a specified amount of time.

Tuning: The act of a receiver being both turned on and receiving a program.

Universe: See *Population*.

Unrestricted Random Sample: A particular type of *probability sample*. An unrestricted random sample is so selected that all possible *samples* of a given size have the same probability of being selected.

Viewing: See *Listening or Viewing*.

Appendix B

COMMENTS OF PRACTITIONERS

This appendix has been included to give the reader an opportunity to review all the opinions submitted by the rating practitioners after they had been given ample time to study a copy of this report prior to publication.

In the interest of easy reading, specific references in their letters to pages and/or paragraphs, in the draft they reviewed, have been altered to conform with the numbering in this printed report, and are indicated by brackets. The typography of this Appendix corresponds to the physical appearance of the original letters.

American Research Bureau, Inc., letter of October 13, 1954:

"Although ARB is not in complete agreement with some of the basic recommendations in this report, notably the definition of television viewing, we certainly feel that the Radio-TV Ratings Review Committee has done an excellent over-all job.

"I should like to state personally that ARB was given every opportunity to submit full data on our operation and to further submit comments on the first draft report. After our comments had been considered, we then received a personal hearing before the entire Committee. In every case ARB was certainly treated with the utmost fairness, for which we express our very great appreciation.

"The decision as to what each basic recommendation should be was, of course, very difficult to make, and the final report can leave no single rating service completely happy. However, in our opinion, the industry could not have selected men better qualified to make these decisions than the members of your Committee, and ARB intends to accept the report and do everything feasible to comply with its specifications."

/s/ James W. Seller

C. E. Hooper, Inc., letter of May 13, 1954:

"The draft of your Committee Report, titled 'Recommended Standards for Radio and Television Program Size Measurements' and dated March 31, 1954, arrived in our office on April 19th. Here are our comments on it, together with some background remarks.

"We wrote to the Advertising Research Foundation on November 6, 1952, seeking assurance '. . . that none of the executive or administrative officers of the ARF is now serving any audience measurement organization in any capacity. . . .'

"It was reported to us in a letter from the ARF dated November 7, 1952 that the president of the Advertising Research Foundation was serving an audience measurement organization. To quote, 'The one exception is Edgar Kobak himself, who, as you know, is a consultant to the Nielsen Company.' Mr. Kobak has been president of the Advertising Research Foundation during the entire period during which this present report has been in the process of preparation.

"Naturally, this situation was deeply disturbing to us. We could not visualize the possibility of Mr. Kobak's disassociating himself from this proposed investigation of the rating methods, and so we could not visualize the possibility of an objective report. That the final report on ratings services was predestined to be what would amount to a promotion piece for the electric recorder method appeared to us to be inevitable from that date forward. However, the report now publishing follows a pattern from beginning to end which exceeds our worst expectations.*

"Below are some specific comments on the report, together with our recommendations as to what should be done with the report.

* * * *

"There are three prime dimensions to which 'standards' can be applied in evaluating audience ratings services. They are:

* ARF asterisk—While ARF does not customarily undertake to refute specific comments, this charge is so unusual and bears so directly upon the integrity of ARF operations that ARF feels bound to set forth the actual facts as they were reported to C. E. Hooper, Inc., in an answering letter dated May 20, 1954. See page 50 for ARF letter.

1. The behavior measured,
2. The audience unit reported on,
3. The sample.

"In spite of its declared objective to judge each service 'operated up to the limit of its practicable potentialities,' the ARF report adopts standards of the lowest, rather than the highest level of significance in two of these three major dimensions:

The standard set for the *behavior* to be measured is proclaimed to be 'tuning.'

The standard set for the *unit* of audience to be measured is proclaimed to be the 'household.'

"It is our considered conviction that these two low standards, 'tuning' behavior and the 'household' unit, have been set with the limitations of the recorder rather than the interests of the broadcast advertiser in mind.

TUNING

"The recorder measures only tuning of radio. A study by Archibald Crossley in 1939 showed that up to 20-25% of tuning was unaccompanied by listening. The difference between the two represents something of no value to the advertiser.

"The limitation on the recorder persists with reference to radio but with added significance now that radio's music program content has expanded several fold. The musical, as contrasted with commercial, portion of radio's material requires and receives less close attention than, for example, dramatic story lines. This means that a listening, rather than tuning, measurement is more to be desired than ever on radio.

"Audience measurement techniques exist, are in use (and were outlined in material submitted to the Committee) which measure radio listening and which can express the difference between listening and tuning behavior on the part of radio audiences.

"Furthermore, the ARF report sets the same low standard, 'tuning,' for both radio and television behavior measurement, ignoring the fact that television's audience behavior includes three measurable levels of attentiveness:

1. Watching
2. Listening, minus watching
3. Tuning, minus listening and watching

"Advertisers have made extensive surveys of 'listening, minus watching' and are aware of the significance of the difference between it and 'watching.' Current figures show 25% 'listening, minus watching' by persons who confirmed this fact by saying 'no one is in the same room with the television set I am listening to' during important morning television programs. A corresponding figure during afternoon programs is 19%.

"A technique exists and is in use which expresses the above three levels of difference in television audience attentiveness. The reporting of '1' and '2' above in a new 'national' television audience reporting service on network television programs is under consideration by one group at this time.

"Though serving an *Advertising* Research Foundation, the Committee overlooks the fact that 'watching' and 'listening,' like advertisement impressions, are psychological experiences. Only the person going through those psychological experiences can testify that they are occurring. They cannot be measured by the movements of a piece of piano wire or by an electric current passing through the wires of a nearby piece of equipment.

"Early in the report the Committee says, in discussing the definition of an 'audience', that there are various levels of attentiveness. 'On the one extreme,' it points out, 'there is a minimum requirement of set tuning. . . . On the other extreme one could require that the exposed individual be given his undivided attention to the broadcast.' Then, shocking as it seems, one paragraph later the Committee recommends that the industry adopt the concept of 'tuning' as its standard of exposure!

THE HOUSEHOLD

"Seemingly motivated by an effort to set a standard low enough to admit the recorder as 'standard,' the Committee has set the 'household' as the unit of audience to be measured. This effort will fail. A 'household' is by definition 'a group of persons living together; a family; pertaining to a family or home.'

"The recorder cannot distinguish between occupied or unoccupied houses. If occupied, it cannot identify the occupants. To admit the recorder, the Committee must lower its 'standard' one more notch, substituting for 'household,' the word *'house.'*

"A technique exists, is in use and has been outlined to the Committee which determines if *people* are attentive.

1. It establishes that the house is occupied.
2. It counts the occupants, if any, who are attentive to programs, and
3. It identifies those attentive by sex and age groups.

"Different members of the household are prospects for different products on different programs. It follows that the unit to be measured should be people, not houses, and that the ambiguity introduced by the word 'household' should be avoided by its elimination. A second aspect of this ambiguity is discussed below.

"The choice of the word 'household', by inference, restricts the place where audience units should be measured to the homes. The measurement of home audiences only represents another limitation of the recorder.

"In television audience measurement, the home may still be an adequate 'place' to measure attentiveness. Like radio sets in radio's early days, the home is now the 'place' where the significant majority of television receivers are located. This may not always be so for television. It is not now true of radio. Radio audiences exist in vast masses in automobiles and in significant totals in retail establishments. In its combined 'tuning-household' standard, the Committee sets the same low standard for radio as for television, denying the radio any measurement of some of its currently most important audience segments.

"A technique exists, is being used and has been outlined to the Committee which does not limit the place where radio's audiences are measured to the 'house'.

THE SAMPLE

"Nowhere in the ARF report is the interviewing, sampling or economic significance of the telephone interviewer's contact with people explored or evaluated except in terms of 20 years ago.

"The percentage of U. S. residences with telephones is now 221% of 1934. Residential telephone subscription stands at 99% saturation in Evanston, 95% in New Rochelle, 90% in Washington, and 88% in the Connecticut factory city where our firm's headquarters are located.

"How different must listening or watching be, in that non-telephone 12%, for the telephone sample not to be representative of *all* homes in Norwalk, Connecticut?

"Apparently, the reason the recorder is given the green light is that it so obviously can measure 'tuning' in 'houses' in that decreasingly significant non-telephone percentage. But so can the Coincidental-Diary method. Why was not the telephone home sample appraised comparatively between radio and television?

RESEARCH BY PROCLAMATION

"The measurement of broadcast audiences is a field where experimentation can be carried out rapidly and easily, albeit expensively. The Committee blithely ignores available empirical evidence already collected, as well as the possibility of collecting new evidence under agreed-upon conditions of scientific control.

"Proclamation of opinions, unsupported by scientific proof of any kind, is not helpful to the industry. There are many equally competent men and women who would be willing to debate the opposite side of most of the Committee's major Proclamations. Neither side could prove its rightness except by experimentation on a meaningful scale. To try to impose these Proclaimed Opinions is to render a disservice to the advertising world.

CONCLUSION

"The ARF report, in our opinion, is one of the most unenlightened and potentially misleading pieces of work to be released by any responsible advertising group in the past 20 years.

RECOMMENDATIONS

"We recommend that the following be done:

1. That ARF take action at once to eliminate from all possible connection with this work any officer or Committee member who has a financial connection with any of the ratings firms whose methods are being evaluated.
2. That all copies, save one, of the current ARF report be destroyed. One should be retained as an example of what not to do.
3. That it undertake experiments, carefully controlled, along the general lines outlined by the Test Survey Committee in 1951 and amplified in our memorandum dated September 11, 1953.
4. That ARF undertake the research necessary to the preparation and publication of a new, fully-documented report based on accepted principles of scientific investigation."

/s/ C. E. Hooper

C. E. Hooper, Inc., P.S. (Dated October 15, 1954)

"Items 3 and 4 on the list of 'Recommendations' which concludes the above letter are appeals by us to the ARF Committee to document its evaluations with original experimental evidence, the Committee having operated without its benefit up to now.

"To demonstrate that significant empirical tests are financially feasible, we conducted a study (such as was recommended) in Los Angeles in August, 1954. We turned over the results of this experiment to the ARF Committee.

"A further purpose was to supply to the Committee data which we believed could not but lead to a re-evaluation of the Coincidental-Diary method (one which we employ) on a significant point on which the Committee has ruled that the method:

'Has telephone and diary biases.'

"In short, another purpose of the test was to demonstrate that our Coincidental 'accuracy control' when applied to diaries removes significant diary biases.

"In the Los Angeles test, therefore, we compared Coincidental-Diary results with Coincidental method results, both measurements being made of identical broadcasts of 406 different programs.

"We used the Coincidental ratings as standard. The authority which qualifies the Coincidental for use as a 'standard' was imparted to it the last time an industry group undertook a methods evaluation. The Coincidental is the only method ever to have been officially endorsed by a tri-partite industry group.

"In its public announcement of its selection and endorsement of the Coincidental method the group gave 'years of experimentation' as the reason it had learned to 'rely' on the Coincidental technique. As indicated in our 'Recommendations', we, too, have confidence in the validity of the 'experimental' approach.

"The industry group which selected the Coincidental method was composed of four advertiser representatives selected by the Association of National Advertisers, four advertising agency representatives selected by the American Association of Advertising Agencies and four radio network Presidents.

"The group was the Board of Governors of the Cooperative Analysis of Broadcasting, Inc. The year was 1945. *To our knowledge no empirical data have ever been brought forward to invalidate the experiment-documented decision of this group.*

"All field work and worksheets were offered for ARF inspection. All the detailed findings from our Los Angeles experiment were submitted to the ARF Committee. Included also were summary tabulations which showed:

I.

"388 of the 406 pairs of ratings, each pair composed of a Coincidental and the corresponding Coincidental-Diary rating of the same program, resemble each other so closely as to fall within the limits expected from two different 400 home Coincidental samples if used to rate the same program. 95.6% of these pairs were observed to fall within these expected deviation limits. The theoretical expectancy is 94%.

II.

"When 246 pairs of ratings were compared by Program Type, the following close approximations were observed:

TYPE (and Number)		AVERAGE RATINGS Coincidental	Coincidental-Diary
Daytime Serial	(14)	1.4	1.5
Children	(36)	5.1	5.3
Sports, Eve.	(12)	6.2	5.8
Variety, Eve.	(79)	6.5	7.1
Drama, Eve.	(44)	7.5	7.5
Mystery, Eve.	(37)	8.6	8.9
Situation Comedy, Eve.	(26)	9.0	9.9

III.

"When all 406 pairs of ratings were compared by rating size groups, the following close approximations were observed:

		AVERAGE RATINGS	
RATING SIZE (and Number)		Coincidental	Coincidental-Diary
Under 10, Day.	(75)	2.2	2.3
Under 5, Eve.	(157)	2.3	2.3
5-9, Eve.	(122)	6.8	6.7
10-19, Eve.	(47)	13.3	13.8
20 & over, Eve.	(5)	27.4	27.6

"The above data, supported by the detailed material submitted to the ARF Committee, can properly be judged to demonstrate that our use of Coincidental 'accuracy control' in our application of the Coincidental-Diary method serves to remove significant diary biases, leaving only telephone bias in the method. Furthermore, original experimental work if conducted by ARF would serve, in turn, to demonstrate that the telephone home sample bias is, in television audience measurement, of comparative insignificance in all but the rarest of instances.

"In short, it leads to but one conclusion, namely, that where significant differences exist between Coincidental-Diary and the industry-accepted Coincidental method they are due to sample size and not due to diary method biases.

"In fact, if the empirical approach were applied comprehensively we believe the Committee would modify its ruling to read:

'Has telephone home sample bias which is observed to be of negligible significance in television audience size measurements in all but exceptional instances.

However, it also provides

'Satisfactory and economical method for producing Station Area projectible and National projectible television audience size measurements.

And in addition provides

'Comparative audience measurements between all types of television homes including non-telephone homes.'

"The Committee rejected our evidence, voting, rather, '. . . to leave the report unchanged'.

"In ruling out our evidence the Committee elaborated on its decision as follows:

> *'Correlation* between any two methods, if established, would not constitute *validation* of *either* method. The results could correlate, and yet one have both telephone and diary "reporting or recording error", and the other only diary or telephone biases of this type.'

"We did not include a 'correlation' study in our exhibits, not because high 'correlation' does not exist but because we judged such to be too superficial for a detailed and important test on a subject of this importance."

/s/ C. E. Hooper

Advertising Research Foundation, Inc., letter of May 20, 1954
To Mr. C. E. Hooper:

"On behalf of the Working Committee on Standards and Methods, which is a sub-committee of the Radio & Television Ratings Review Committee of the Advertising Research Foundation, we wish to acknowledge receipt of your letter of May 13th.

"We appreciate the comments, observations and criticisms you have made therein on the technical matters involved. We have also received observations from other research organizations to whom we submitted the draft of the Committee's report. They will all be thoroughly considered and taken into account in the further deliberations of the Committee.

"At this time, we should like to answer the comments you have made concerning Mr. Kobak. They are not well-founded and we are certain that you would want to know the true situation.

"Mr. Kobak attended the first meeting of the full Radio & Television Ratings Review Committee on July 28, 1952. He addressed the Committee on the importance of the matters it would be dealing with. He also disclosed to the Committee that he numbered amongst his own personal clients users of rating services and a rating service as well. He thereupon left the meeting and has never since attended any meeting of the full Committee or of any of its sub-committees. Neither has he ever expressed by written or verbal communication of any sort, to any person on the Committee or any of its sub-committees, or in any way connected with the work of any of them, his views as to anything they were doing or proposed doing. We on the Committee have no idea of what he personally thinks of our work and he has never in any manner, directly or indirectly, attempted to influence the course of our deliberations or activities on any subject."

/s/ G. Maxwell Ule
Chairman of Working Committee
on Standards and Methods

/s/ Donald W. Coyle
Member of Working Committee
on Standards and Methods

/s/ E. L. Deckinger
Chairman of Radio & Television Ratings Review
Committee and member of Working Committee
on Standards and Methods

C. E. Hooper, Inc., letter of June 8, 1954:

"In my letter of *May 13th*, I confined my technical remarks to: 'Tuning,' to the 'Household' and to a limited reference to 'Sample.' Because your letter of *May 20th* says *'They will all be thoroughly considered and taken into account in the further deliberations of the Committee,'* I wish to call your attention to other technical points which we also consider to be of importance. Remarks are confined almost entirely to what appears in Chapter [III] and on one page in Chapter [II], because we believe those pages will get the major attention of the industry if the report should be released. We do not, in short, agree with everything else appearing in the report. This present communication also represents no change in our previous conviction that no part of the report, as it now stands, should be released.

"In Chapter [II], page [14], under 'Methods' appears 'The Telephone Recall.' I am sure these words will connote, to your readers, either 'overnight recall' or 'day-part recall.' Each of these methods was used, at different times by the Cooperative Analysis of Broadcasting.

"When those same words appear in the descriptive title: 'The Telephone Coincidental and The Telephone Recall,' I believe they will carry that same connotation, which is incorrect. A title which will carry correct connotations could read: 'The Telephone Coincidental and Immediate (15 Minute) Recall,' if it is designed to describe the method of which we are listed as a 'Practitioner.' If this revised description is acceptable to you, it should be substituted everywhere in the report.

"Also in Chapter [II], page [16], C. E. Hooper, Inc. is omitted as a 'Practitioner' using 'The Telephone Coincidental' method.

"We are enclosing specimen reports published by us during the past nine months using 'The Telephone Coincidental Method' in a continuing service. (See Exhibit 'A').

"Also in this section 'The Personal Coincidental Method' is incorrectly described in the 'Practitioners' column as being used in 'No continuing service.' C. E. Hooper, Inc. uses it in both radio car-audience surveys and retail-store audience radio surveys. Samples of a recent survey of each type is shown in Exhibit 'B'. C. E. Hooper, Inc., accordingly, should appear as a 'Practitioner' of the 'Personal Coincidental Method.'

"Because the 'Personal Coincidental' does have a 'Practitioner,' it should appear as a method treated in Chapter [III], beginning page [17]. It should also be credited with measurement of those segments of the 'out-of-home' audience referred to above.

"Also in Chapter [III], beginning page [17] under 'Basic Information Standards' Item 5, 'The Measurement Should Be Representative of All Households' carries a 'No' for both 'The Telephone Coincidental' and 'The Telephone Coincidental and (revised) Immediate Recall.' We can assure the Committee that investigation, on this point, will demonstrate that there is no statistically significant difference in the 'local' audience size estimates obtained from the telephone sample: (1) on 'popular' programs (2) on low-rated programs of general appeal (3) on many programs of specialized appeal, and (4) on practically all programs in most of the areas where we are conducting local measurements and in which residential telephone subscription is above 80%. Differences in audience behavior, in the 20% non-telephone homes, must in short be tremendous as compared with the telephone-home sample for the 20% factor to upset the audience pattern established in the 80% segment of the all-home sample.

"In Chapter [III], on [second pull-out page] under 22C, the report credits 'The Recorder' with 'Yes' indicating no reporting errors other than 'mechanical failures.' Investigation by the Committee comparing the significance of the tapes *received*, in time for tabulation, (compared with the tapes placed in the *theoretical* recorder sample) will, we believe, lead to a further qualification in terms of *retrieved* tapes. The latter, not the former, is the 'recorder' sample and it varies from report to report. We believe such inspection is indicated because of the major descriptive heading appearing on this Section, namely, 'The Potential for Radio-TV Audience Size Measurements Under Highest Standards of Research Implementation Based on Present Known Technology.' The difference between the placed and the retrieved tapes is a 'known' difference having a 'technological' effect on the audience size estimate. To hold tabulation and publication of the report until all tapes were received, would be technically impracticable.

"In this same section, the 'Telephone Coincidental' gets a 'No'. The reason given (namely, 'problems of incomplete reporting to household mem-

bers') is, we believe, not based on empirical data, rather it is based on propaganda distributed in recent years by persons trying to justify the higher ratings on radio produced by 'Aided Recall.' If this is so, we believe you will agree that it has no place in this report unless first substantiated by the Committee with experimental data. Furthermore, there is no distinction drawn in the report between radio and television on this important point.

"In Chapter [III], on Page [21] under Item 12, we believe the relatively great significance of 'Audience Characteristics' and the relative insignificance of 'Household Characteristics,' calls for a different treatment than that followed in the report.

"We suggest either literal adherence to 'yes' vs 'no' description of 'Audience Characteristics' (without reference to 'Household Characteristics') or the addition of another item 'Household Characteristics' on which a number of methods will qualify in addition to 'Recorder.'

"Also, on the subject of 'Audience Characteristics', no specific sample size standards are set for such measurements. In referring to local samples, the figure of 400 households is specified for the 'audience size estimates.' We see none for 'Audience Characteristics.' On a program, rating 10.0 in a sample of 400 households, the audience lending itself to analysis of 'Characteristics' is 40. Is that adequate? The report does not say."

/s/ C. E. Hooper

A. C. Nielsen Company letter of May 12, 1954:

"We appreciate this opportunity to comment on the ARF report 'Recommended Standards for Radio and Television Program Audience.' Art Nielsen is in Europe at this time, thus preventing his careful study of this important report. In his absence we have endeavored to give this report the same meticulous consideration which he most certainly would have given this final draft.

"Since this report deals with research methods of securing audience measurements, it is natural for us to be greatly interested in the soundness and the quality of such an appraisal. We are greatly appreciative of the commendable job done by the committee working on this project, and we feel that the committee's realistic attitude on cost and practical point of view regarding the value and importance of reliable methods is the surest way by which higher standards for all audience measurements can be achieved.

"We do have one major concern regarding this report. It is clearly stated and understood by the ARF committee writing this report that this is *strictly an evaluation* of the *potential* quality possible by each different method if the highest degree of refinement is used in carrying out the method in practice. There is a danger that many in the radio and television industry will associate this appraisal with the commercial organization using the method regardless of the degree to which they actually carry out many of the important refinements indicated as possible. This is tantamount to assuming that each concern has endeavored and succeeded to the same degree in achieving maximum refinements.

"We realize that the committee would not recommend such an assumption, because they have clearly expressed the need for a study of the degree to which commercial operations achieve these standards. Many of the deviations from ideal practices are of great importance and create cost differentials most advantageous to those who least carry out the committee's recommended standards.

"For this reason we hope every effort will be made to minimize the chance of such misuse or misinterpretation of this ARF report.

"In reviewing the report, we understandably found many instances where differences in points of view existed. However, in getting down to spe-

cific comments, we recognize and respect the committee's important objective of having the report published as promptly as practicable.

"Therefore, rather than take issue with all of the conclusions with which we cannot wholly agree, we simply offer for the committee's consideration the following additional suggestions:

1. In Paragraph [three, on Page 28], the statement '. . . households are "interviewed" several times, the statistical equivalent of these requirements should be provided' could be expanded slightly to '. . . households are "interviewed" or measured several times, the statistical equivalents of the samples used to meet these requirements should be provided. Similar statistical equivalents are possible when the trend accuracy is evaluated from several "interviews" or measurements of the same households.'

   The above change is suggested because of our conviction that more adequate attention should be directed to the error reductions achieved through the use of fixed samples.

2. On Page [6] in Section B9 of Chapter [I], reference is made to 'Unduplicated household audience to two or more broadcasts.' We believe that this would be clearer if split into two separate and distinct 'supplementary information standards.'

   (a). Cumulative audience—the number of homes reached by one or more broadcasts of a program over a series of broadcasts, i.e., one-week, two weeks, four weeks.

   (b). Audience duplication—the number of different homes reached by one or more programs and/or spot schedules used jointly (radio, tv, or both).

3. On Page [8], in the last paragraph of Section [I], 11, it is stated that share of audience cannot be computed for network programs, because of varying times of broadcast in different areas. It is recognized that it is not possible in all cases to ascertain share of audience, but, in our experience, this can be done whenever the preponderant portion of the program facilities in a large geographical area (such as a time zone) is used at a single time of broadcast. Many programs use facilities in this manner, one time of broadcast predominating in the East and Central Time Zones, another on

the Coast. Under these circumstances useful share comparisons can be produced.

"We hope you will extend to each member of the committee who has participated in this ARF report our view that it represents an important contribution toward better appreciation and fuller understanding of factors determining the usefulness of audience measurement information. Recognition of these factors creates healthy environment for progress in audience measurement research.

"With kindest personal regards,"

/s/ C. G. Shaw

The Pulse Incorporated letter of May 6, 1954:

"Enclosed is our comment for inclusion in your appendix to the report."

/s/ Sydney Roslow

"We are of the opinion that what will be printed as 'Appendix [B]' should actually be in a foreword to the Report, 'Recommended Standards for Radio & Television Program Audience Size Measurements.' Delays in publication, attended by gossip and surmise, combined with considerable leakage of the first draft in the past months, already have caused considerable uneasiness among prospects. Undoubtedly the mark-time situation has penalized other rating services—perhaps some more than Pulse, enjoying steady expansion. But we are conscious of unsigned subscribers, because services following the same profit-motive which guides agencies and advertisers can not afford the luxury of ruminative leisure for theorizing, as can mutual non-profit organizations. Hence from the practical and pragmatic point of view, we hope that readers will scrutinize the Appendix of your report which treads heavily on the personal or corporate body's most sensitive nerve, the pocketbook.

"In the preface to your report, under 'Presentation of Results' (i.e., the very heart and core of the subject) it is stated that, 'This part of the report is thus a compromise between a wholly theoretical and a practical viewpoint.' And again it is stated that 'what is desirable may, at least at our present level of knowledge, be unattainable at *reasonable cost*.'

"One of the services publicizes the fact that it has lost $6,000,000 on its electronic recorder. Yet not to be cutely remote—your analysis boils down to an excellent promotion piece for the A. C. Nielsen audimeter, if we may by-pass double-talk.

"Since you give Nielsen the benefit of an asterisk on the opening page of Section [II]—kindly toss one in for Pulse, too, and please note:

> As one of the services stretched out on the Procrustean bed of the 'wholly theoretical,' it should be called to the attention of the readers of this report that Pulse, has in operation *its* all-electronic recorder, named DAX, which requires no tapes, no supporting diary to reduce expense. Pulse's device is not a 'Goldberg' either. DAX has no lights, buzzers or sirens. And it is 'instantaneous.'

"The 'Instantaneous' concept, a good ad-slogan, seems to be new since the general discussion of last year—and two years, almost, since Pulse devoted its annual luncheon to a discussion and support of ARF objectives. That concept does not seem to be altogether practical.

"Bearing in mind that we know the electronic recorder's 'instantaneous' impression first-hand from investing thousands of dollars to develop the prototype DAX you can examine in our office—we are not satisfied with the pull-out pages that tick off the 'Yeses' and 'No's' a little too handily. Pulse, remember, *pioneered* in measuring out-of-home listening. It is the sole service that *can* and *does* make regular studies of out-of-home listening for radio! That battery and portable sets and other listening swelled audiences by 20% for the past two years should not be casually paired against a 'Yes for Television.' In our estimation, piddling-size machine samples, including what our DAX can produce at a common-sense price—not at a loss of millions, we assure you!—can not stand up against the validity and superiority of our carefully pre-tested U.S. PULSE TV, which provides:

67,000 interviews a month;
25,000 interview measurements of each daytime program;
42,000 interview measurements of nighttime programs;
6,000 interview measurements per nighttime program;

"These measurements, observe, are in considerable excess of the 1,200—acceptable minimum you ask for in your report.

"Further, in our experience we have found that people *do* have the capacity to identify correctly their viewing and listening, as substantiated in tests by Pulse and others. For example, 97% of families do *not* pretend to have been listening to radio when they were *not;* and the rightness of a mass of critical, even hostile checks and data, confirm Pulse's time-line association that precludes mis-use of printed roster, with a degree of accuracy that has made Pulse the #1 service in point of subscriptions. Commercial usage, as you indicate, is the measure of utility at a practical price—and our service is not theoretical but in operation *now*. Your committee's observations about memory-failure, prestige bias, and incomplete reporting 'black-wash' most of the services, your 'whitewash' of the electronic recorder being the main exception.

"Actually, in probability sampling, is the substitution of a household in

the placement of a small panel, electronically metered, the answer to *non-response?* Isn't the refusal to accept a meter for this tiny sample more serious than the refusal of a personal interview, in a huge national sample? And as for market-by-market meters and counters with diaries —it seems to us that 68 households provide very thin statistical ice to skate on!

"Again, isn't mechanical failure, the refusal of the machine to answer or express itself with clarity a plain form of non-response? How negligible is this factor—the recorder's mechanical failures with 10% to 20% tape spoilage acknowledged! And can the mechanical failures be separated from non-response—or eliminated by adjustment—or made up for by diaries, as contemplated?

"In Section [I], Page [5] a delightful non-sequitur is stated: 'the committee recognizes the value of a total audience measurement, but designates it as a supplementary standard.' With that jolly little denial of your principal objective—why then all the bother?

"We should like to make a final observation:

"From the beginning we have objected to your committee's neat 'Yes-No' check-lists, which we conscientiously pasted into pull-out sheets. By the time those are given the convincing substantiality of print, and widely broadcast (the 190 members-plus) and widely distributed, including loan to numerous account executives, and 'peeks' to others—you are giving sound rating service objectives a black eye. Which in turn gives other research a black eye!

"Time and usage settled so many problems. For long years the Metric and English standards have lived peaceably side by side. For a long time the various rating services have lived peaceably side by side—except when an occasional zealot has crusaded a little too violently. We have a reasonable belief in the rightness of personal interviewing, which does not have to compromise with the practical, and which does not have to use cost-cutting substitutes (which ironically wind up by being far more expensive on a subscription basis!).

"Without resorting to picayune criticisms, please look at the observation on Page [28], Section [IV]—and the very startling conclusion: 'To get the same reliability in local audience size measurements, the same sample size (1,200) would be required. *However, less precision is re-*

quired for local measurements.' In the old days that line of reasoning used to be called specious, if not downright meretricious—and if you remember your high school Latin, that last is derived from 'meretrix.' We wonder how the 'spot' advertiser feels about your conclusion!

"We call attention to the fact that when accurate information is wanted, responsible organizations do not stake important decisions on recorders, telephones, diaries, or adjustment formulae. *They go out and talk with people:* Who does that? Government services, U.S. Census, Pure Food & Drugs, U.S. Department of Commerce & Labor; U.S. Army, Air Force, Navy; Congressional investigations; F.B.I.; all insurance companies; National Better Business Bureau; finance and loan company look-ups; readership studies and services, such as Daniel Starch for magazines; Look's studies by Crossley; Life and other studies by Politz; Crowell-Collier studies by Simmons; Fortune studies by Roper; public opinion studies by Gallup & Robinson, Roper, Scripps-Howard, Hearst and many others; Dun & Bradstreet financial reports; police department and detective bureaus' investigations; private industry studies by the thousands, thousands conducted by advertising agencies; newspaper polls, magazine polls (and do not forget the literal killing of the Literary Digest by the famous poll of questionnaires sent by mail to telephone subscribers and automotive license lists); Nielsen's own 100,000 coverage report.

"Research is both dynamic and creative—and even when theoretically perfect not always practical. Remember the story of Alfred Sloan and his group who were searching for the perfect solvent? 'Hmm—it would dissolve everything, you say?' commented a farmer they were talking with. Assured that was correct, that the perfect solvent would dissolve anything, the farmer inquired, 'What would you keep it in?' We question that any one service is able to serve all users. What ARF blandly recommends in committee for radio and television is a monotheistic worship of the machine, and a monopolistic concentration on a one-and-only service. Even if it existed—who of the committee members' agencies or clients could afford it?"

# # #

Trendex, Inc., letter of May 6, 1954:

"Attached is the ARF Committee's Report on 'Recommended Standards for Radio and Television Program Audience Size Measurements.'

"Again, we think the committee has done a splendid job of presenting its recommendations. There is, however, one basic contention which the Committee makes with which we heartily disagree.

"In Chapter [I], page [2], under, 'Information Standards' the Committee recommends that the Industry adopt the concept of *tuning* as its standard of exposure, because it is the most objective of the various levels of exposure . . . that all other measurements are more subjective. The Committee states that tuning is the only measurement that is *'unambiguous.'*

"We feel that the Committee has a case as far as the objectivity of its recommended measurement goes . . . tuning is tuning, period. *But,* the Committee makes an ambiguous statement when it recommends tuning as the standard for the Industry for the basic audience size measurement, because a measurement of sets 'tuned' to a program is *not* a measurement of the size of the program's audience.

"Isn't it much more *unambiguous* to accept the word of a respondent as to whether he or anyone in his home was or was not part of the program's audience? In all other phases of research . . . readership studies, audience reaction studies, depth interviewing, door to door surveys, the word of the respondent is the base for final evaluation.

"If we accept tuning as our standard audience size measurement, we are approaching the antiquated circulation technique used by newspapers and magazines to sell space. Tuning is a mechanical process. The device now used to record tuning is a machine . . . but audiences are not machines, and they will never be mechanized.

"We strongly urge the Committee to reconsider its decision in this matter."

/s/ Robert B. Rogers

Videodex, Inc., letter of May 17, 1954:

"I have attached the following;

(A) DRAFT: 'Recommended Standards For Radio and Television Program Audience Size Measurements' Dated March 31, 1954.

(B) VIDEODEX Comments on the above; 7 pages

"We appreciate the week extension given us on submitting our comments. I understand that several of the other services took advantage of this additional time.

"I only wish we had more time to expand our comments. This draft came to us in the midst of a very large special study we were doing and also during a time when several members of our company were travelling around the country. The three weeks given us was short enough by itself and with the additional items above, did press close. However, I know we all have busy schedules, but it is unfortunate that a project as important to us all as this one should be penalized to any extent by a stringent time dimension.

"Please contact me if I can be of additional assistance.

"Best personal regards."

/s/ Allan V. Jay

VIDEODEX COMMENTS

"THE REPORT STATES . . . 'For unrestricted random samples your Committee recommends a minimum sample size of . . . 1,200 households in final tabulation for each national program audience size measurement.' [pg. 28]

OUR COMMENTS ARE: A 1200 minimum national sample, bearing any semblance to allocation by television set density or population would find less than half the present markets with one or less member of the sample. This would become a multi-city statistic rather than one nationally projectable. Even the properties of disproportionate sampling would not accommodate the problem with this size sample.

Furthermore, an unrestricted sample of a random nature, is practically unattainable. By definition, this type of sample implies that all units have an *equal* chance of selection. The more attainable goal would be a probability sample where all units have a *known,* as opposed to an *equal,* chance of selection. The number of cases should be considered in terms of such a construction. In all practical cases some stratification is involved.

"THE REPORT STATES . . . 'since attended sets cannot be defined unambiguously, tuning is recommended as the standard for the Industry.' [pg. 2]

OUR COMMENTS ARE: The text should cite a difference between TV and AM. TV sets turned on have more likelihood of being attended in the physical sense of presence than AM. The Committee ridicules the distance from the set criterion as a source of ambiguity. This ridicule does not constitute refutation per se. For example, one person might be relatively hard of hearing and sit near to hear; another person might enjoy sitting beyond ten feet away. These do not constitute ambiguous situations in themselves. They are inherent characteristics of viewers and listeners. (Please refer to comments on [pg. 3] and [pg. 4])

"THE REPORT STATES . . . 'A program audience size measurement reported on a household base has wider application than a measurement based upon individuals.' [pg. 3]

OUR COMMENTS ARE: The household base is not as meaningful from a marketing evaluation standpoint as viewers. Again, the committee's aims should be high. The question could be settled simply as to which is the most important to the advertiser-; viz,— Do TV sets buy products or do people? If the former, then the committee's selection of the household base is supported; If the latter, then it should be rejected and the 'viewer' concept substituted.

"THE REPORT STATES . . . 'Your committee recommends the average instantaneous audience size measurement as the standard.' [pg. 4]

OUR COMMENTS ARE: The average instantaneous audience measurement does not satisfy the ideal requirements for true audience measurement and effective program exposure. It is analogous to counting a person who touches three magazines at a newsstand and decides to buy one as a reader of all three. This attempt at sensitivity measurements is a desirable objective by the committee insofar as it relates to a standard program unit in television and radio; e.g. 15 minute. The *Average Total Audience* type of measurement satisfies this and has the property of fuller projectability, which average instantaneous audience does not.

"THE REPORT STATES . . . '. . . cumulative number of households which have tuned to any part, over some specified minimum, of a program broadcast.' [pg. 6]

OUR COMMENTS ARE: The committee has within its power to set a minimum for proper program credit that would aid the industry considerably; i.e. by setting a standard of the viewing in minutes required for inclusion in the audience. This decision need merely require a consensus of committee member experts and other members of the industry, and would be no more arbitrary than other standards agreed on by the committee.

"*THE REPORT* STATES . . . 'Your committee, therefore, recommends as a supplementary measurement a rating for each program as a per cent of all households. For national audience size measurements, this percentage would be based on all households in the United States.' [pg. 6]

OUR COMMENTS ARE: Per cent of all households would not necessarily reflect meaningful popularity levels, even as a supplementary measure. One program, of a network originating type, might clear stations in 40 markets. This limited clearance might be due to the inability of the network to clear time and/or might match the distribution objective of the advertiser's product. Another show might be in 150 markets. Both are network shows by definition included in a national rating tabulation. Any program percentage based on the respective networks of either show and applied against all households in the United States would have little meaning, even in the supplementary sense.

"*THE REPORT* STATES . . . 'For this supplementary standard, therefore, your committee recommends that an 'attended sets' concept be applied.' [pg. 9]

OUR COMMENTS ARE: 'Attended Sets' should be a primary standard and not a supplementary standard. Please refer to [pg. 3].

"*THE REPORT* STATES . . . 'For 'strip' programs, your committee recommends that the audience size measurement be based upon the average of all of the broadcasts of that program during the week of the measurement.' [pg. 9]

OUR COMMENTS ARE: This destroys the detail needed for participating sponsorships. Individual daily segments may have different adjacencies and competing programs. The intra-week fluctuation in ratings and sets in use would not be observable under such a standard. The cost of television is such that all advertisers cannot afford to buy the entire strip and have to settle for one day, and others are forced to settle for part of each day. They may want to study sustaining segments to augment or serve as a substitute for their present unit of the same strip.

"THE REPORT STATES . . . '. . . measurement should be available for the total program or customary segments. These will usually be for 15 minutes or multiples thereof.' [pg. 10]

OUR COMMENTS ARE: Please refer to comments on [pg. 4]. This applies to sustaining as well as commercial segments and embraces the need for a 15 minute practical standard which is compatible with the average total audience concept of reporting.

"THE REPORT STATES . . . 'The A. C. Nielsen Company recently announced its plans for syndicating a local audience size measurement service which is a combination of the recorder and diary methods. Because this method was not being practiced commercially, at the time this report was written, it will not be discussed.' [pg. 14] footnote

OUR COMMENTS ARE: However the implication of Nielsen's reversion to a diary panel as a framework for local reporting has had a decided effect on the committee's preparation for this present report. Further, the Nielsen local service has budget commitments from buyers with a tentative starting date later this year. This is therefore an immediate and practical consideration and should be included in this report.

The reason given by the committee for not including it at this time since it is not commercially practiced is inconsistent, since three other methods—(1) personal coincidental (2) personal unaided recall (3) telephone recall, are included in this report and yet have no practitioners at this point nor even the remotest prospect of such in the near future.

Nielsen's technology is descriptively known to the market. It does represent a different method from those covered. Buyers are actively entertaining purchase. If the committee is to fulfill its function and allow the ARF as a group to keep faith with its subscribers then there is a clear duty to include the new Nielsen local service. Failure to do so on the reasons advanced [pg. 14 footnote] would only succeed in weakening the value of the report to the industry.

"THE REPORT STATES . . . 'Of the fundamental methods, the recorder, the personal coincidental and the telephone coincidental can deliver an average instantaneous audience estimate. For all practical purposes, all other fundamental methods can only deliver a total audience measurement, although theoretically, an open end (In open-end diaries, the respondent could be asked to record to the nearest minute tune-in and tune-out for each program) diary could produce an approximation to an average instantaneous audience.

The combination methods can do it through adjustment by the coincidental part of the measurement.' [pg. 20]

OUR COMMENTS ARE: The phrase, 'although theoretically' should be taken out since the entire report is premised on statement found in [pg. xvi]; . . . 'This is your committee's judgment of the various methods reviewed in terms of their theoretical promise under optimum implementation.' Chapter [III] is clearly entitled ' . . . MEASUREMENT METHODS.'

FURTHER COMMENTS: 'Combination methods can do it through adjustment.' It is a known fact that the biases peculiar to a single method when combined with another separate and distinct method are compounded in combination. There are analogies to this situation. For example, a car can take me anywhere in the world. Of course when I come to water, I go by boat. The combination of boat and car allow me to go to my objective. Again, yellow and blue are my favorite colors. When they are combined they no longer have the original identifiable properties. Given more time by the committee it would be possible to gather statistical examples from the social sciences and related fields on this phenomenon.

I can combine anything I'm doing with something else and attain the missing element at each stage. This encouragement by the committee in recognizing 'combination techniques' as desirable can and will be carried to the extreme by services lacking so-called 'fundamental' features as prescribed by the committee.

"THE REPORT STATES . . . 'Audience size measurements based upon characteristics of the individuals exposed to a program, such as age and sex, can be provided by all methods except the recorder. This is because only the recorder is restricted to measuring whether the set is turned on or off.' [pg. 21]

OUR COMMENTS ARE: Please refer to our comments on [pg. 3], regarding the fundamental value of 'attended sets' standard as opposed to household tuning. We maintain the position that a more meaningful evaluation of 'audience worth' could be realized if the committee's ranking of primary and supplementary measures were reversed concerning household tuning and attended sets.

"THE REPORT STATES . . . 'However, a probability sample design applied to a fixed panel tends to become biased through time because of drop-outs and changes in the composition of the universe covered. A fixed panel operation must therefore develop some technique for maintaining probability design through time.' [pg. 22]

OUR COMMENTS ARE: Explanation and distinction should be made between a fixed panel concept and a rotating sample or moving sample. The electronic recorder panel is relatively fixed and is inherently inflexible. The diary panel is not a panel by design in this sense, but more a rotating sample. Panel implies a static quality. Since the text refers to a panel or diary panel the implication is strong that a rotating sample and a fixed panel are synonymous.

"THE REPORT: At [pg. 23] following #20, there should be another category entitled—'The Measurement Should Be Able To Measure Trends Meaningfully.'

In a panel *type* of operation the rating variations from one reporting period to another is due primarily to actual audience shifts, whereas changes in rating levels from all other types of samples is greatly due to the statistical variation in successively new samples being taken for each reporting period.

"THE REPORT STATES . . . '. . . No—in a practical sense only a total audience measurement is possible.' [first pull-out]

OUR COMMENTS ARE: Please refer to our comment on [pg. 4] and also [pg. 10]. Also comments on [pg. 20], first part.

"THE REPORT STATES . . . 'Unduplicated audience (household) to two or more broadcasts; Diary—Yes' [first pull-out]—

OUR COMMENTS ARE: Only if a panel is used. Please refer to [pg. 20], paragraph [#4].

"THE REPORT STATES . . . 'Your committee specifies that the minimum samples should be (1200) households (national). Even for television audience size measurements, it is not necessary that all the cases be television homes.' [pg. 28]

OUR COMMENTS ARE: Please refer to our comments on [pg. 28]. The meaning of the second sentence cited above is not clear to us and might be less clear to readers of the report. Therefore, it might be wise to spend another sentence or two in the final report in elaborating further.

"THE APPENDIX STATES . . . *'Full Network Audience.* The audience of all network stations during a given time period regardless of what program is carried on any one station during the given time period.' [pg. 36]

OUR COMMENTS ARE: Please refer to our comments on [pg. 28]. It is our position that the above referred to measure cannot be obtained meaningfully through a national sample of 1200, which would find less than half the network station markets with one or less member of the sample. Given time by the committee we would be able to give examples of this point."

/s/ Videodex, Inc.

ARF NOTE

As a result of the published service comments the following changes were made in the report:

1. On page 19, modified the statement regarding the representativeness of telephone samples.
2. On page 16, added a footnote regarding C. E. Hooper's and The Pulse, Inc., contemplated recorder measuring method.
3. On first pull-out, altered the chart to conform with the text discussion of representativeness of telephone samples.
4. On first pull-out, split "Audience Characteristics" into two categories.
5. On page 20, elaborated on the footnote.
6. On page 28, fifth complete paragraph, elaborated on explanation of sample sizes.