

How to Conduct Consumer and Opinion Research

THE SAMPLING SURVEY
IN OPERATION

EDITED BY
ALBERT B. BLANKENSHIP, PH.D.
Managing Director
National Analysts, Inc.



Published under the sponsorship of
AMERICAN COUNCIL ON PUBLIC RELATIONS
SAN FRANCISCO, CALIFORNIA

HARPER & BROTHERS PUBLISHERS
NEW YORK AND LONDON

How to Conduct Consumer and Opinion Research

Copyright, 1946, by Harper & Brothers
Printed in the United States of America

All rights in this book are reserved.
No part of the book may be reproduced in any
manner whatsoever without written permission
except in the case of brief quotations embodied
in critical articles and reviews. For information
address Harper & Brothers

THE AMERICAN COUNCIL SERIES OF PUBLIC RELATIONS BOOKS
Rex F. Harlow, *Editor*

Contributors

WROE ALDERSON—A.B., George Washington, 1927. United States Department of Commerce, 1925-1934; market analyst, Curtis Publishing Co., 1936-1943; partner, Wroe Alderson, Simon & Sessions, 1944-. Author of numerous articles and studies on domestic and international economic marketing problems. Member, American Marketing Association, Institute of World Economics.

GEORGE H. ALLEN—M.B.A., Harvard, 1938. Assistant to president, National Theatre Supply Co., 1938-1940. Assistant, Research and Promotion Department, WOR, New York, 1940-1942; research manager, WOR, 1942-1944; assistant director of promotion, WOR, 1943-1944. Manager and secretary of Board of Governors, Cooperative Analysis of Broadcasting, 1944-. Member, American Marketing Association, American Statistical Association. Consultant, United States Treasury Department. President, New York chapter of American Marketing Association. Numerous radio articles.

SAMUEL G. BARTON—Investigator, Market Research Interviewer and Field Supervisor, Special Service Department, Retail Credit Co., 1932-1938; partner, Industrial Surveys Co., 1938-; president, Industrial Surveys Co., 1942-; member, American Marketing Association. Author of radio and research articles.

ALBERT B. BLANKENSHIP—Ph.D., Columbia, 1940. Director of research, Hartwell, Jobson & Kibbet, 1939-1940; Rockefeller Fellow, 1940-1941; director of market research, N. W. Ayer & Son, Inc., 1941-1944; managing director, National Analysts, Inc., 1944. Author: *Consumer and Opinion Research*, Harper & Brothers, 1943. Author of many papers on research technique. Member, American Marketing Association, American Psychological Association. Fellow, American Association for Applied Psychology. Member, Society of the Sigma Xi. Lecturer, Temple University, 1945-.

ALBERT A. CAMPBELL—Ph.D., Stanford University, 1936. Assistant professor of psychology, 1936-1942, Northwestern University. Fellow, Social Science Research Council, 1939-1940. Assistant head, Division of Program Surveys, Bureau of Agricultural Economics, 1942-. Member of American Psychological Association, Society for the Psychological Study of Social Issues, Sigma Xi. Author of articles in the field of learning.

CORNELIUS DUBOIS—B.A., Harvard, 1926; promotion manager of *Time*, 1931-1937; research director, *Life*, 1937-. On leave as deputy chief, Bureau of Intelligence, Office of Facts and Figures and OWI, 1944. Member of Advisory Committee for WPB Surveys of Civilian Requirements. Member, American Marketing Association.

MARJORIE FISKE—M.A., Columbia, 1938. Research and administrative assistant, Princeton Office of Radio Research, 1937-1938; research associate, Market Analysts, Inc., 1938-1939; assistant director, Field Service Division, National Federation of Business and Professional Women, 1939-1941; research director, National Federation of Business and Professional Women, 1941-1943; senior associate, Bureau of Applied Social Research, Columbia University, 1943-. Author of articles on research techniques. Associate, American Psychological Association.

- ALBERT D. FREIBERG**—Ph.D., Cornell, 1936. Various assistant, associate director of Market Research Division, director of Market Research Division, and vice-president, Psychological Corporation, 1935-. Member, Market Research Council, American Psychological Association, Copy Research Council, Society of the Sigma Xi. Fellow, American Association for Applied Psychology. Lecturer, Columbia University, 1939-.
- MORRIS H. HANSEN**—M.A., American University, 1940. Statistical assistant to director of United States Bureau of the Census; currently president of the Washington Chapter of American Statistical Association and a Fellow of the Institute of Mathematical Statistics. Publications include a monograph and several articles, principally on sampling and the application of sampling in the fields of social and economic statistics.
- CHESTER E. HARING**—A.B., Western Reserve University, 1908; Columbia School of Architecture, 1909. Vice-president and director of marketing, Batten, Barton, Durstine & Osborn, New York, 1923-1934; director, Market Research Division, Hooper-Holmes Bureau, Inc., New York, 1936-1942; Committee for Economic Development, New York. Director of Marketing Research, 1944-1945. Harvard-Bok Advertising Award for research study for American Radiator Company, 1925; Harvard Bok Advertising Award for research study for Johnson & Johnson, 1927; Annual Advertising Awards—Medal Award for research for Haring Index of Local Business Conditions. Harper & Brothers published *The Manufacturer and His Outlets*, 1939. Member, American Marketing Association.
- PHILIP M. HAUSER**—Ph.D., University of Chicago, 1938. Assistant director of United States Bureau of the Census; vice-president of American Statistical Association and Secretary of the Population Association of America. Formerly on Sociology faculty of University of Chicago. Publications include several monographs and a number of articles in the field of population and social statistics.
- ERNEST R. HILGARD**—Ph.D., Yale University, 1930. Successively, assistant professor, associate professor, professor of psychology, and head of department, Stanford University, 1933-; on leave as assistant director, Civilian Relations Division, Office of Civilian Requirements. Member, American Psychological Association.
- C. E. HOOPER**—M.B.A., Harvard Graduate School of Business Administration, 1923; president, Clark-Hooper, Inc., 1934-1938; president, C. E. Hooper, Inc., since 1938. Member of American Marketing Association and Market Research Council. Coauthor, *Radio Audience Measurement*, Stephen Daye, Inc., 1944.
- DANIEL KATZ**—Ph.D., Syracuse, 1928. Various instructor, assistant professor, and associate professor, Psychology Department, Princeton, 1928-1942. Associate director, Reports Section, Division of Program Surveys, Department of Agriculture, 1942-1943; variously director of Field Staff, research director, and deputy director, Surveys Division, Office of War Information, 1943-1944. Chairman, Department of Psychology, Brooklyn College, 1944-. Coauthor, *Social Psychology*, 1938. Member, American Psychological Association, Society for Psychological Study of Social Issues; Secretary-Treasurer of Society for Psychological Study of Social Issues, 1944-1945. Author of numerous articles.
- NELLE KELLY**—M.A., Pennsylvania, 1941. Field supervisor, Department of Commercial Research, Curtis Publishing Co., 1942-1943; assistant director, The Pulse, Inc., 1943-.
- ARTHUR KOLSTAD**—Ph.D., Columbia University, 1933; associate director of the Bureau of Research, 1925-1929 in the Western Washington State Normal School; director of research of the Houser Associates, Inc., 1933-1938; partner in the Houser Associates, 1938; Social Science Analyst in the Information and Education Division of the Army Service Forces directing the preparation of reports on studies of attitudes of civilian employees of the War Department, 1944.

- PAUL F. LAZARSFELD—Ph.D., Vienna, 1925. Director, Division of Applied Psychology, University of Vienna, 1927-1933. Rockefeller Fellow, 1933-1935; director, Office of Radio Research, 1935-. Director, Bureau of Applied Social Research, Columbia University, 1944-. Associate professor of sociology, Columbia University, 1943-. Author: *Radio and the Printed Page*, Duell, Sloan and Pearce, 1940; *Radio Research*, 1941, Duell, Sloan and Pearce; *Radio Research*, 1942-1943, Duell, Sloan and Pearce, 1944; and *The People's Choice*, Duell, Sloan and Pearce, 1945. Author of numerous technical articles. Member, American Marketing Association, American Psychological Association, American Sociological Association, American Statistical Association; Fellow, American Association for Applied Psychology. Consultant to Office of War Information, War Department, War Production Board.
- A. W. LEHMAN—New York University, 19—. Assistant managing director, Association of National Advertisers, 1934-1941; general manager and secretary, Cooperative Analysis of Broadcasting, 1934-1941; technical director, Advertising Research Foundation, 1941-1944; managing director, Advertising Research Foundation, 1944-; president, Cooperative Analysis of Broadcasting, 1944-. Author of numerous articles on radio. Member, American Marketing Association.
- LAWRENCE C. LOCKLEY—Ph.D., 1931, Harvard. Associate in English, University of California, at Los Angeles, 1921-1927; Member, Research Staff, Harvard Business School, 1929-1930; professor of marketing and head of department, Temple University, 1930-1935; technical expert, Bureau of the Census, 1935; market analyst, Division of Commercial Research, Curtis Publishing Co., 1935-1943; manager, Market Research Division, E.I. du Pont de Nemours & Co., Inc., 1943-. Author of volumes on letter writing; author of numerous articles on research techniques. Member, American Marketing Association.
- HERBERT C. LUDEKE—B.S., 1923, M.I.T., Assistant to the vice-president in charge of sales, Eaton, Crane and Pike, Pittsfield, Mass., 1928-1929; in charge of research, Boston office of Curtis Publishing Co., 1930-1939; currently manager of Development Division, Curtis Publishing Co., Philadelphia. Member, American Marketing Association; national director and vice-chairman of the New England chapter of the American Marketing Society, just prior to its absorption by the American Marketing Association. Coauthored one paper which appeared in *Sociometry* entitled, "A Technique for Validating Interviewing Methods in Reader Research." Author of several articles for trade papers.
- CHARLES E. NOYES—A.M., Brown University, 1931. Director, Civilian Relations Division, Office of Civilian Requirements, War Production Board, 1943-1944; assistant and associate director, Division of Information, WPB, 1942-1943. Author, *Economic Freedom*, published by Harper & Brothers, 1943.
- SYDNEY ROSLOW—Ph.D., New York, 1935. Research associate, Psychological Corporation, 1933-1940; director, The Pulse, Inc., 1940-. Member, American Marketing Association, American Psychological Association; Associate, American Association for Applied Psychology.
- ALFRED C. WELCH—Ph.D., Minnesota, 1938; research manager, Knox Reeves Advertising, Inc., 1939-1942; assistant to the director, Services Programs Division, Office of Civilian Supply, WPB, 1942-1943; assistant to the vice-chairman of Civilian Requirements, WPB, 1943; chief, Special Surveys Branch, Civilian Relations Division, Office of Civilian Requirements, WPB, 1944; market director, Knox Reeves, Advertising, Inc., 1944-. Associate American Psychological Association. Member, American Marketing Association.
- MATILDA WHITE—M.A., Radcliffe, 1937. Boston representative, Percival White, Inc., 1933; research assistant, Department of Sociology, Harvard, 1933; principal executive, Market Research Co. of America, 1937-. Chief consultant to Office of Civilian Requirements, War Production Board, and Survey Division, Office of War Information, 1943-. Author of books and articles. Member, American Marketing Association.

PERCIVAL WHITE—A.M., Harvard, 1910; managing director of Percival White Engineering Works (England) 1912-1914; manager, Percival White Engineering, 1915-1917; automobile expert to Ordnance Dept., United States Army, Rock Island Arsenal, Ill., 1917; director of development and foreign representative in Europe, Aluminum Mfrs., Inc. (Aluminum Co. of America), 1919-1920; general manager, White Engineering Works, 1943; president, Percival White, Inc., Marketing Counselors; president, Market Research Corporation of America; managing director, American Economic Foundation, 1935. Member, President Hoover's Emergency Committee for Employment, 1930-1931, Market Research Council, American Society Mechanical Engineers, Society Automotive Engineers, Society Industrial Engineers, American Marketing Association, American Statistical Association. Author: *Market Analysis*, 1921; *Motor Transportation of Merchandise and Passengers*, 1923; *Forecasting, Planning and Budgeting in Business Management*, 1926; *Business Management*, 1926; *Advertising Research*, 1926; *Scientific Marketing Management*, 1927; *Sales Quotas*, 1929; *How to Fly an Airplane*, 1930; *Marketing Research Technique*, 1931.

WALLACE H. WULFECK—Ph.D., Yale, 1938. Associate director of marketing research, Psychological Corporation, 1939-1942. Vice-president in charge of research and personnel, Federal Advertising Agency, 1942-. Member, American Marketing Association, American Psychological Association; Fellow, American Association for Applied Psychology.

Contents

I. Introduction	<i>A. B. Blankenship</i>	I
PART ONE. SURVEYS REPORTED TO BUSINESS AND INDUSTRY		
II. Market Description—Quantitative and Qualitative	<i>Lawrence C. Lockley</i>	11
III. Research for Product Development	<i>Percival White and Matilda White</i>	29
IV. Making a Technical Industrial Survey	<i>Chester E. Haring</i>	43
V. The Movement of Branded Goods to the Consumer	<i>Samuel G. Barton</i>	58
VI. Psychological Brand Barometer	<i>Albert D. Freisberg</i>	71
VII. Public Relations Research	<i>Wallace H. Wulfeck</i>	86
VIII. Consumer Research in the Development of Advertising Copy	<i>Alfred C. Welch</i>	96
IX. Copy Testing	<i>Albert D. Freisberg</i>	118

X. The Office of Radio Research <i>Paul F. Lazarsfeld and Marjorie Fiske</i>	137
XI. The New Cooperative Analysis of Broadcasting <i>A. W. Lehman and George H. Allen</i>	151
XII. The Coincidental Method of Measuring Radio Audience Size <i>C. E. Hooper</i>	156
XIII. The Personal Interview-Roster Method of Radio Measurement and Its Application <i>Sydney Roslow and Nelle Kelly</i>	172
XIV. The Radio Panel <i>Samuel G. Barton</i>	181
XV. The Survey Applied to Editorial Problems <i>H. C. Ludeke</i>	188
XVI. How Many Readers? <i>Cornelius DuBois</i>	209
XVII. Measurement of Employees' Attitudes and Morale <i>Arthur Kolstad</i>	222
XVIII. The Last Three Feet <i>Chester E. Haring</i>	232
PART TWO. SURVEYS REPORTED TO GOVERNMENT	
XIX. The Surveys Division of OWI: Governmental Use of Research for Informational Problems <i>Daniel Katz</i>	241
XX. Sample Surveys in Census Work <i>Philip M. Hauser and Morris H. Hansen</i>	251

	CONTENTS	xi
XXI.	Surveys of Consumer Requirements <i>Charles E. Noyes and Ernest R. Hilgard</i>	259
XXII.	Attitude Surveying in the Department of Agriculture <i>Albert A. Campbell</i>	274
	PART THREE. SURVEYS REPORTED TO THE PUBLIC	
XXIII.	Trends in Public Opinion Research <i>Wroe Alderson</i>	289

HOW TO CONDUCT
CONSUMER AND OPINION
RESEARCH

Chapter I

INTRODUCTION

ALBERT B. BLANKENSHIP

National Analysts, Inc.

RECENTLY there has been increased use of the questionnaire survey technique.¹ Opinion polls are coming in for increasing attention. Congress has investigated the polling methods. The so-called commercial surveys are being conducted for all types of business and in connection with many different kinds of problems. Government is making extensive use of the approach in respect to many postwar problems.

Immediately following the collapse of Germany and the Japanese surrender, the use of surveys increased tremendously. Many companies returning to the development of consumers' goods were faced with problems about what products to make, how to price them, whether they could be sold, where they could be sold, and how they could be sold. The survey approach is the only adequate means of answering many of these problems.

Because there is already a great demand for survey work, probably the greatest demand in the years that the approach has been used, a number of companies and individuals now advertise themselves as "research firms" or consultants but actually have little or no background in the work. For this reason it is especially important now for the businessman who needs research to be able to select the *right* man or organization. That is one of the purposes of the present volume: to provide a sufficient understanding of the methods used in applying the sampling questionnaire technique to different phases of activity so that the businessman will be able to determine whether his choice of personnel for the work is a wise one. At the same time, the present volume

¹For purposes of this book, we are defining the questionnaire survey as any study using a set list of queries to obtain replies from a "cross section" of people. The theory underlying the so-called "sampling survey" is that a representative group of people will give replies that will resemble closely the replies of the total population from which the sample has been selected. Discussions of this theory and methods for making it work in practice can be found in other sources.

2 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

shows what organizations offer continuing (and fairly standardized) measurement services, so that the businessman will be able to get, from one source, the answer to his question on the various services available.

As time goes on, there is little question that specific training for the field of questionnaire sampling will be available. Right now few colleges offer a well-rounded program of courses for such training; the individual specializes in psychology, sociology, advertising, economics or statistics, and then rounds out his education in the techniques *after* his formal education has been completed, by getting practical experience in the field. With the development of specialized training programs for the field, there will be an increase in specialized training courses for the questionnaire surveyor. This volume, representing the cumulative research background and original thinking of some twenty competent men, should be a useful text in such courses. Naturally, use of this volume in such a way presupposes elementary knowledge of techniques, since these are certainly an essential for the person going into the field. On the other hand, the book has intentionally been so written that the businessman without such background can gain knowledge of the services available in the field, and the various types of application that are possible with the procedure.

Lastly, the book should be useful to the "part-time" research man. For example, in some of the smaller advertising agencies the research function is combined with some other activity, and there is no chance or reason for the man in the position to become a real research specialist. This volume ought to be useful in acquainting him with the various kinds of survey available, and should make more apparent to him the possible values that the approach could have in connection with the accounts of his organization.

FOR WHOM ARE SURVEYS MADE?

In general, surveys are conducted for the benefit of three principal groups of people: commerce and industry, the government, and the public.

1. *Commercial surveys.* Surveys conducted for business and industry fall logically into several categories: those of products and markets, those of advertising and selling, and those of company personnel.

Description of the market is important for companies that know where neither the present nor the potential market for a particular product lies. A company contemplating production of an electric kitchen mixer should have some idea of the present and potential market in terms of such factors as geographical section, economic level, and age groups, in order to plan how to market the product. Dr. Lockley's chapter explains how the sampling survey can be used here.

Again, if a company is contemplating the sale of a new product, such as a dehydrated tomato soup, it is necessary not only to perfect the product technically but to make sure that people like the product. The chapter by Percival and Matilda White explains in some detail their approach to this problem. Mr. Haring's chapter "Making a Technical Industrial Survey" illustrates a phase of this technique as applied to industrial equipment.

Mr. Barton and Dr. Freiberg explain the use of their organizations' syndicated services on the movement of branded goods. Such reports help to tell a manufacturer who is buying his goods, when and where customers are gained and lost, and in general help to give the producer a current picture of the movement of his items versus competitive items.

The purchase of branded goods is, of course, a measure of the degree to which the public believes in a particular company. At best, this is an indirect measure; people may still buy Brand X even though they dislike the company, merely because that happens to be the best brand on the market. Dr. Wulfbeck explains how the survey can be designed to measure public opinion toward a particular company, a group of companies, or an industry.

Among studies concerned with advertising and selling, Dr. Welch's article is concerned with the development of effective advertising: the means by which surveys can be utilized to learn what should be said and how to say it in advertising. Dr. Freiberg's exposition of copy testing, which includes a discussion of the syndicated services offered for magazine and newspaper advertising, concerns measurement of the effectiveness of advertising that has already been published.

Treatment of the material on measurement of media has been limited to radio and magazines, largely because these are the two fields that have come in for most survey attention. There were necessarily some five chapters on radio alone, because of the many organizations offering surveys as a means for measuring radio and its effectiveness. Dr. Lazarsfeld and Miss Fiske have written a notable chapter that not only serves to tell the reader the different kinds of measurement available, but explains how certain specialized studies may offer more than measurement of audience size alone. This chapter is followed by those on the syndicated radio audience measurement services: C. E. Hooper, Cooperative Analysis of Broadcasting, The Pulse, and the radio panel of the Industrial Surveys Company.

Measurement of the readership of publications can best be discussed in two steps: measurement of readership of competitive publications, and measurement of reading within a single publication. Mr. DuBois explains adequately the various techniques available in the competitive measurement, and tells about the continuing studies of *Life* readership. Mr. Ludeke discusses the methods that can be used for measuring what has been read within

4 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

a particular publication, and goes into some detail to show how this has been applied to the *Country Gentleman*.

The chapter on measurement of employee morale, by Dr. Kolstad, is the first of two concerned with studies among company personnel. Dr. Kolstad explains how the questionnaire method can be used to measure morale, and demonstrates how the results can be applied within an organization. Mr. Haring then takes over with his material on dealer studies, and tells the method used in a survey among gasoline dealers.

2. *Government surveys.* Unfortunately, because of government restrictions, some of the most interesting survey work done by or for the government cannot be revealed. Fortunately, enough of a story can be told about how the government makes use of these studies for the imaginative to assume how the army made use of surveys in the Quartermaster Corps, and in measurement of military morale. In addition, applications of public opinion measurement by the Overseas Branch of the Office of War Information are not too hard to visualize without any chapter on the subject.

Dr. Katz has done an excellent job in explaining use of sampling surveys by the Domestic Branch of the Office of War Information. Dr. Hauser and Mr. Hansen show admirably the various ways in which the bureau of the Census has used sample studies in place of complete censuses. Dr. Hilgard and Mr. Noyes explain how the War Production Board has made surveys of consumer requirements, and Dr. Campbell explains the survey operation of the Department of Agriculture.

3. *Public opinion polls.* In the polling operation, the basic technique of sampling surveys is similar to those used by business and government; in fact it is the opinion poll most commonly known by the public, since it is these results which the average citizen has access to. Mr. Alderson shows the operating methods of both candidate and issue polls, and describes the Gallup, Crossley, *Fortune*, and National Opinion Research Center operations. Alderson's analysis of the predictive accuracy of the major polls in the 1944 presidential election is especially interesting and timely.

THE TYPES OF SURVEYS

Survey organizations can be classified according to whether they offer regular services, irregular though specialized, or generalized service.

1. *Regular, syndicated services.* In the field of commercial surveys there are a number of organizations which offer continuing, periodic studies. Reading of printed advertising, for instance, is measured by services of Grant and Starch. Periodic measurement of radio audiences is reported by the Cooperative Analysis of Broadcasting, Hooper, The Pulse, and the Industrial Surveys

Company. Movement of branded goods is measured periodically by Industrial Surveys Company and the Psychological Corporation. The *Life* study offers a continuous measurement of the audience size of the weekly magazines. While the reports are not available for public consumption, such magazines as *Reader's Digest*, *Coronet*, and the *Saturday Evening Post* all continuously measure the items within a particular issue that have been read.

The government surveys are likely to be least regular; most government studies are conducted to answer a specific problem when it arises.

In the field of the opinion polls, Gallup regularly syndicates both election and issue polls through newspapers. The *Fortune* poll results are reported regularly in that magazine.

2. *Specialized services.* A number of organizations which use the sampling questionnaire method do not offer regular, continuing studies, but have, nevertheless, specialized to a noticeable degree. Wroe Alderson & Company specialize in the field of industrial research. Willmark Research Corporation performs most of its work among dealers. Industrial Surveys Company does most of its research by means of the "panel" technique, in which the same group of people continuously answers questions. Audience Research Corporation specializes in measuring reactions to moving pictures. C. E. Hooper, in addition to its continuing radio audience measurement, handles only problems dealing with radio. Houser Associates concentrate their efforts in the measurement of industrial morale.

This degree of specialization varies with the organization. The Pulse, for instance, while primarily interested in the measurement of radio audiences, will take on almost any sort of questionnaire survey. The Psychological Corporation, while it developed and sells its Brand Barometer, will also handle almost any sort of sample study. National Analysts, Inc., although doing continuous readership research for various publications, will also undertake practically any kind of questionnaire survey.

3. *"General practitioners."* In addition to the research organizations which have, to one degree or another, specialized, there are in the field many firms that have not developed any specialty at all, but who are competent to take on almost any sort of work. Such organizations believe, rightly, that the survey procedure is a basic technique which, with proper application, can answer problems in many fields. This is intended as no reflection on the specialists in the field, who also can do a competent job but prefer to concentrate their efforts in one particular segment of the work.

PLAN AND PURPOSE OF THE BOOK

1. *Plan.* The basic outline of the book has already been covered under the section "For Whom Are Surveys Made?" This seems to require no expansion

at this point; the various writers have handled that portion of the job competently.

It does seem in order to point out the method that most of the writers used in preparing their material. They all started with the same assumption: that the reader knew little about the specific techniques used in survey work. Fortunately, Dr. Lockley's chapter, the first, discusses the basic techniques sufficiently so that later references to terms or methods should readily be understood by those who have first absorbed this chapter.

Each writer was asked to follow a rough sort of outline in preparing his material. He was requested first to tell about the general and specific purposes of studies in his particular field. He then was asked to outline some of the basic methods available in survey techniques for securing answers to the problems. Finally, he was requested to cite a particular study, giving its principal points of procedure, its major results and conclusions, and any action that was taken on the basis of the recommendations.

The degree to which each writer has followed this outline has, fortunately, varied. Some contributors are considerably more interested in the first portion of coverage—the general discussion. Others spend more time on their discussion of a particular project. This makes for a well-rounded volume, even though it does mean in a few cases that discussion of the general place of surveys in a particular field is light.

2. *Purpose.* This balance between discussion of applications and discussion of particular projects should make the book valuable to all three classes of people for whom it was intended. Business people should be able to get enough of both purposes and cases to meet their requirements. Students should find this an acceptable case book interspersed with discussion of purposes. Part-time researchers or beginners in the field ought to be able to pick up a reasonably good idea of the purposes and methods of surveys so that their orientation to the field will be aided.

One point needs clarification. This is a book on the questionnaire method. As such, it intentionally does not cover some adequate methods for getting answers to some of the problems outlined by the contributors. For instance, the Nielsen Index, a continuing inventory of retail outlets for food and drug items, is a reliable measure of brand movement, but since it does not involve the asking of questions, there is *no discussion of it in this volume.*

In other words, the questionnaire method, discussed throughout this volume, is only *one* means of obtaining information on the problems outlined. In the case of business and industrial problems, there are literally reams of priceless statistical data collected by the Census Bureau and by private organizations. Since these are not *sample studies*, they are not discussed here.

Again, in the field of media, a valuable measure of comparative worth is

total circulation of publications. But this, too, is not a questionnaire or a sample study, so has not been included.

One minor exception to this rule of inclusion only of questionnaire studies has been permitted. When the questionnaire method has been combined with another technique, obviously it would be unfair to present only the questionnaire portion of the material.

PART ONE

SURVEYS REPORTED TO
BUSINESS AND INDUSTRY

Chapter II

MARKET DESCRIPTION—
QUANTITATIVE AND QUALITATIVE

LAWRENCE C. LOCKLEY
E. I. duPont de Nemours & Co., Inc.

THE shortest route from factory to market for any product is along the line of the desires and the convenience of the consumer. But the consumer is a multiple personality who lives in Coronado, Calif., in Millinocket, Maine, and in all the cities, towns, hamlets, and farmsteads in the 3,700 miles between. It is the place of the market research technician to find out which groups of the millions of consumers living between Coronado and Millinocket can advantageously buy his wares, how much they are likely to buy, where they would like to buy them, how much they will be willing to pay, and in what form they will want the wares to be.

Market research is an ancillary service that allows management to manufacture with a better idea of what can be sold, and how much can be sold, and to market with a better idea of how to combine the various tools of selling to minimize waste. It is, therefore, a subservient science. It is the responsibility of management to match the information provided by market research with other information on the availability of raw materials, power, transportation, labor, and price relationships, and to draw conclusions that will maximize the social efficiency of industry. While market research can help management, it should not attempt to usurp the functions or responsibilities of management.

Just what are the limits of the service that market research can offer management? In a field as fluid and as rapidly expanding as this, it is difficult to establish fixed boundaries. As is indicated by other chapters in this book, the techniques of market research are susceptible to many applications. But from the standpoint of market determination, there are eight commonly employed

uses—uses, by the way, equally applicable to industrial and to consumer marketing problems.

1. *Testing new products.* When new products are developed, it is common practice to submit them to the type of consumer for whom they are intended for appraisal or opinion. Such consumer tests may require only a short interview—as when a new perfume is being made ready for the market or they may entail elaborate performance tests. At all events, such tests give management a helpful appraisal of the suitability of the new product for its market, and often provide a comparison between the new product and the products with which it will have to compete that will guide the establishment of marketing policies for it.

2. *Appraising the attitude of consumers toward established products.* Because new competition is usually encouraged by every successful product on the market, and because consumers' needs and desires are in a constant state of flux, no manufacturer should rest on his oars when one of his products has attained wide acceptance. It is desirable for him to find out how well satisfied his customers are with his wares, how completely these wares fit consumer needs, whether his market is expanding with the growth of generations, or whether it will decline in popularity when the present generation of users retires from the market.

3. *Determining the strength of competitive products.* It is just as important for a manufacturer to know what his customers or his prospects think of competitive products as it is to know what they think of his own. It was no decrease in quality of product that made the Model T Ford car obsolete—it was an improvement in the quality of competitive cars.

4. *Measuring the extent of the market for products.* It is the job of the market research man to be able to estimate within reasonably close tolerances how large a market there is for products. This job is probably the most difficult one in the repertoire of the market analyst because the sales of the particular product depend only in part on the sales policies and activities of its proprietor. Improvements in competitive offering or increased sales efficiency on the part of competitors may decrease the share of the market that might be reasonably expected. It is, therefore, not a job that can be done and forgotten; it must be continually in progress.

Moreover, for many new products, an actual estimate of the market the product will enjoy is impossible, because it is impossible to forecast the rapidity with which consumer buying habits can be changed. The market research man can calculate with considerable accuracy the dollar volume of breakfast foods the American public is likely to consume during a particular year. But even with the benefit of consumer tests of a new and distinctive breakfast food as a guide, it is impossible to predetermine the share of the market that this

new breakfast food will have. Only the most tentative estimates can conscientiously be made.

5. *Locating new uses for products.* Frequently conditions of consumption change or a product reaches the practical limits of its market development. If the market for such a product is to grow, it can do so most handily through the stimulation of new uses. A prominent brand of cleansing tissue was sold in substantially greater volume (as were its competitors) when new uses were located and stressed in advertising. The summer slump in the consumption of coffee was decreased when it was found that consumers could be persuaded to drink iced coffee.

This aspect of market research is particularly important currently, because a large number of new products and new machines have been developed during the war, for war uses. Many of these products or machines will pass into disuse unless market research can find new uses for them. In this instance, it is important that laboratory research and market research join forces. It is rare that the market research man is able to locate new uses, particularly for technical products and materials, without some guidance as to the limits of application inherent in them.

6. *Determining channels of marketing.* It is easy to say that the convenience of the purchaser and the efficiency of the marketing functionaries should determine how products are to be marketed. Actually, it is not so easy to determine how the purchaser's convenience can best be suited, nor whether low sales volumes are due to the inappropriateness or inefficiency of the distributors, the jobbers, or the retailers. One firm, long marketing an electrical product through one set of wholesalers, introduced a new product associated with the first. Sales were disappointing, until a market analyst reported that these wholesalers were not logical suppliers for the retailers who should handle the new product.

Should a product be given restricted distribution through a few retailers in each market? Or should the manufacturer try to get as many retailers as possible to stock the product? Should he sell through factory branches to retailers or employ distributors and wholesalers? There are, of course, some classes of merchandise for which trade practices are so solidly established that there can be no doubt as to channels of marketing. But in many fields the answers to such questions can rarely be given without careful market research.

7. *Checking the efficiency of distributors, wholesalers, and retailers.* Even when what appears to be the most satisfactory combination of marketing channels has been determined, it is desirable to audit the system periodically, to be sure that it is kept to a high pitch of efficiency. In fact, the audit should not stop with the middlemen through whose hands a product passes, but should include the sales activity of the manufacturer as well. Many firms have

14 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

failed to maintain the share of market they might have enjoyed because their marketing machine developed inefficiencies.

8. *Determining unsatisfied needs of consumers.* One of the most fruitful applications of market research technique is the attempt to find out what consumers—whether individual or industrial—do not have that they want. Rarely are such wants formulated in the minds of any class of consumers. But a careful appraisal of what people have, how they use what they have, and what they think of what they have will usually show an opportunity for new products. One firm with large research facilities undertook to find out how well satisfied housewives were with their kitchen duties. Each housewife recounted to the interviewer every task she performed in her kitchen throughout a day, and then answered the question, for each task, "Was that hard to do, or did it take too long?" A promising market opportunity for a number of new products grew out of this study.

These eight applications of market research techniques are not, of course, all the ways in which the market analyst can be helpful to management, but they do indicate the directions in which market research has been most frequently used.

To summarize, the task of the market research technician is, first, to find out what can be marketed and, second, how it can best be marketed. But these are very general questions. In order to approach his task efficiently, he must approach his job by systematic steps. Possibly the following divisions will adequately describe the planning and execution of a market survey to determine the qualitative or quantitative aspects of a market:

1. Statement of the problem.
2. Formulation of the questionnaire.
3. Selection of the sample.
4. Field supervision.
5. Editing and tabulating the questionnaires.
6. Analyzing the results.

STATEMENT OF THE PROBLEM

At first, it may seem strange that the most important step in market research is the statement of the problem, since research projects are rarely requested if there is not a problem. Yet probably more research budgets are wasted through failure to formulate the specific problem than through any other cause. It would be difficult, for instance, to know what to do to satisfy a request for "a study of the market for mechanical refrigerators." By implication, every aspect of marketing and of consumer demand must be treated. On the other hand, a survey that attempts to find out the per cent of current sales of mechan-

ical refrigerators which are replacement sales is definite, and can be organized.

It may well be that a major survey must include the consideration of a number of different questions. But unless each question is written down in specific form, with a high degree of definiteness, the survey will be difficult to handle.

Sometimes it is difficult, without resorting to field work, to formulate the problem which a survey is expected to solve. Often all management can be sure of is that a particular product is not selling at a profit, or that selling costs for one of a company's lines are too high, in comparison with other items, or sales for a particular item are falling off. It may be desirable for the research man to talk with a few dozen customers informally merely to freshen his point of view and to rid himself of the "office point of view."

It is a common experience for the research worker to find that the problem posed him by management is not actually the problem that management wants solved. A survey may be requested to analyze the deficiencies of retail distribution for a particular product; yet the decreases in sales may be due to a change in consumer wants or to increased sales efficiency of the marketing efforts put behind competitive products. Probably the isolation and formulation of a problem to be studied is as important a contribution to management as any market research can make.

It is for this reason that the market research executive should work with sufficient flexibility to change the direction of his work as the problem on which he is working emerges.

WHAT KIND OF INFORMATION WILL HELP MANAGEMENT?

At the least, it is the responsibility of the market research man to supply information that will help management in deciding policies and making decisions. But the necessary information may come from any one of a number of sources. The source most frequently overlooked, but often most valuable, is the sales or trade records of the individual firm. Many marketing questions can be answered by a careful analysis of these figures—if they are available and have been maintained in usable form. Questions on satisfactoriness of sales territories, justness of sales quotas, profitable size of order, the need for additional wholesale or retail representation, and the profitableness of marginal accounts can often be answered largely from such a study. Moreover, an appraisal of individual items in a line of products can often be made, and the need for changes in product design or for a readjustment in price relationships can be discovered early. Other types of information gathering and analysis may be needed in conjunction with the analysis of a company's own records. But the utility of this reservoir of marketing information should never be overlooked.

A large body of market information is available in the many publications of federal, state, and local governments and in the publications of trade associations. This general classification of data may sometimes be difficult to apply to particular *marketing problems*, because it is necessarily general, whereas the information needed must be specific. The market researcher who wants to know whether consumers prefer paint in quart or two-quart containers cannot find his answer in government statistics. But he can find out a great deal about the over-all consumption of paint in the United States and about the channels of distribution used by *manufacturers of paint*. In many cases, the specific information must be located by the individual researcher, and then can be generalized with the help of the over-all figures available from public sources.

The mainstay of market research is now, and will probably long remain, information obtained by questioning many people in various localities as to their past market behavior and their attitudes. In general, there are four ways of obtaining information from people—whether they are consumers, retailers, wholesalers, or any others who influence the marketing process. The research technician may ask questions by mail of a group with whom no previous arrangement for answering has been made—the familiar old mail questionnaire. He may arrange with a panel or group of people (presumably a well-selected cross section of the total group he is interested in) to send in regular reports of some or of all the purchases they make or to answer whatever questions he may send to them. In most cases, panel operators compensate the co-operating group in one way or another. He may send interviewers or go himself to talk with individuals and to ask them questions. Finally, he may observe them as they buy, as they walk past a window display, or as they do anything that reflects their buying habits and attitudes.

Each of these methods of obtaining information from people is appropriate to some problems. It is rare that they may be regarded as alternate methods or that they should be put in competition with each other. And each has its advantages and disadvantages.

THE MAIL QUESTIONNAIRE

The mail questionnaire is a useful resource of the market research worker—but one which is in some disrepute because it has been misused. As a method of measuring the market it is faulty, in that there is no method of sample control. The research man never knows whether the people who answer a mail questionnaire are or are not typical of his total market. There is usually a tendency for people most interested in a product or service or problem to answer mail questionnaires about it. Rarely are these people representative of the total group, and generalizations made about a total group on the basis of

the relationship between favorable and unfavorable answers from those most interested are likely to be misleading.

On the other hand, the mail questionnaire is an entirely adequate device for obtaining many kinds of market information. The adequacy of advertising copy, the effectiveness of sales appeals, the sources of satisfaction or dissatisfaction, and other qualitative facts may be learned by means of the mail questionnaire, if it is possible to phrase questions with enough clarity and simplicity to make them readily understood without oral explanation and to allow easy rapid answering.

There are definite problems in the use of the mail questionnaire. The first has to do with the mailing list to be used. When a rather compact and easily identified group of people is to be questioned, it is usually possible to obtain a mailing list covering the group. If, on the other hand, the group to be questioned is to be a random sample of the population, the problem becomes more difficult, because often a mailing list must be built from heterogeneous sources, and because it is difficult to get names and addresses of people representative of the lower income groups.

The second problem is that of cost. If all costs—direct and overhead—are counted, it usually costs around 10 cents a name to mail a questionnaire. If a return of 10 per cent is reached, the cost per response comes to an even dollar—an amount usually adequate to cover the cost of personal interviewing. The mail questionnaire offers no economy, as compared with personal interviews, unless the percentage of returns is higher than average or unless the group to be questioned is thinly scattered over a wide geographical area.

The third problem is that of generalizing the results. Although the mail questionnaire is not a satisfactory instrument for *measuring* market conditions, it is often used for this purpose. Even though a majority of the respondents who return questionnaires are enthusiastic in their approval of a particular commodity, the fact that a small minority of the people receiving questionnaires bothered to answer them suggests the probability that those who did answer have not been typical of the entire group. The market research man, therefore, has no basis for generalizing the returns to describe the entire group questioned. It is only when the proportion of returns climbs well over the 50 per cent level that a convincing case can be made for the mail questionnaire as a measuring device.

THE CONSUMER PANEL

The consumer panel is a means of obtaining information from consumers which has had a rapid growth in popularity. It is a well-selected sample of whatever group the researcher desires to study which has promised to send in regular reports on certain classes of purchases or to answer questionnaires

18 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

mailed from time to time. As a usual practice, members of a panel are compensated for each reply by a stipulated number of "points," which may be accumulated and exchanged for premiums. When the members of a panel are organized around some idea or theme, their interest may often be maintained for considerable lengths of time. When Dr. Franklin R. Cowl organized a panel among farm families for the *Farm Journal*, he developed a "Home Makers' Guild" and aroused the interest of farm families in keeping track of their family expenditures. This device appeared to arouse a degree of interest more active than the premiums supplied for co-operation.

The greatest value of a panel appears to lie in its allowing regular reports from the same group of families, so that the flow of purchases or the flow of consumption of particular types of products can be studied over a considerable period of time. How accurate a measuring device the consumer panel will be depends on the satisfactoriness of the sample, and the skill with which interest and co-operation are maintained by the operator of the panel. Two questions with respect to the use of consumer panels have not yet been satisfactorily answered: first, whether panel members soon become self-conscious in their purchases and thus become atypical of the total group they are supposed to represent; second, what rate of turnover is desirable. People die, move away, lose interest, change their ways of life through marriage, change of occupation, bereavement, etc. It is clear that an absolutely stable panel would, after a period of time, cease to be a good sample of the total group to be studied. And it is probably true that a panel with a very high rate of turnover would be ineffect in co-operation. It is difficult to set any figure which would indicate a desirable rate of turnover.

The consumer panel is an exceedingly valuable tool of market research—but it is not a cheap tool to operate. The costs of building the panel are high. When it is decided what characteristics must be included to constitute adequate representativeness, then families in all parts of the country must be selected to meet these characteristics, investigated to make sure that they qualify, invited to membership—preferably by personal invitation on the part of an interviewer—and then trained to co-operate. When families drop out, they must be replaced by other families of like characteristics and probably in the same geographical location. The cost of premiums is a considerable item. All in all, consumer panels are among the most expensive methods of obtaining information on the American consumer.

PERSONAL INTERVIEWS

The conventional way of obtaining market information from consumers is by means of the interviewer, who goes from house to house with a question-

naire. Properly used, the interviewer method will give sound guidance through authentic reports from a well-selected sample of the consuming public or of any part of it which needs to be studied. *Perfunctorily used, this method can be virtually useless.* There are several principles which should guide the market research worker in using the consumer-interview method of obtaining information.

The first consideration is the proper selection of interviewers. The validity of the entire survey rests on the alertness and integrity of the interviewers who serve as the connecting link between the inarticulate consumer and the market research technician. It is important, therefore, that these interviewers are faithful in their work and intelligent in their interpretation of their assignments. In most of the larger cities, a relatively small group of women have drifted into consumer interviewing, have gradually come to be on the lists of interviewers of a number of firms, and may, in a month, interview for a dozen different employers. Many of them have never had any training in interviewing and have never met representatives of the firms for which they have worked. And quite a few of them have become routine and perfunctory about doing work they do not understand.

If the market research worker is able to go into the field himself, or have a competent assistant do so, and select new interviewing material, he will usually find these newer interviewers more responsive to training and direction. *Probably the most important part of the training of interviewers is the inculcation of a professional pride and the realization of the economic and social utility of market research work.* This "indoctrination" will do more than anything else to carry an interviewer through the more discouraging phases of her (for most consumer interviewers are women) work. It gives her a way of answering the stock objection, "Why is that question any of your business?"

It is probably a mistake to try to get all interviewers to work in the same way. *A woman whose personality is particularly winsome may be able to get the information she asks for on the strength of this personality, without the help of any carefully prepared "approaches."* Whatever the personality or bent of the individual interviewer, she will do better work if she figures out her own methods of approach and her own ways of getting the friendly co-operation of the people she interviews. Her training should emphasize the need for getting uncolored and complete information, for avoiding asking questions in a leading manner, for meticulous following of sampling instructions, and for a thoroughness in interviewing that minimizes that booby trap for tabulators, "No answer."

On-the-spot supervision is always better than giving interviewers their instructions by mail. The supervisor who is able to meet interviewers before

they start out on a survey, go over instructions with them, and answer the questions that are certain to be present in the minds of the interviewers can turn in a better job. Moreover, he can check interviews each day, go over them with the *individual interviewers at the end of each day to check for completeness and accuracy*, and can see to it that sampling instructions are followed. He can even accumulate interviews for a day or two, and then have his interviewers exchange work and try to do a little informal field tabulation. This device does more to improve the quality of interviewers' work than any other.

All this makes for good, accurate, complete, and penetrating field work. Unfortunately, it is rarely possible within the limits of most commercial research budgets. When a few interviews from each of many localities must be obtained to build up the desired sample, the time and money costs of on-the-spot supervision are too great. If interviewers can be given personal supervision every few months or at least once a year, and if they are good interviewers, they can usually be depended on to turn in a good job if they can be induced to read their instructions carefully and to write or wire for clarification on any points not entirely clear to them. One research firm, Wroe Alderson, Simon & Sessions, has developed a "quiz" to go with each set of instructions to interviewers. *These questions can be answered only after a careful reading of the instructions.*

It is always good practice to send back for correction any interview forms not completely or clearly filled out. All too often, time schedules make it impossible to send forms back for correction in time for use in the tabulations planned. But even if the amended form is returned too late for use, the interviewer is impressed with the fact that her work is actually scrutinized and is likely to improve her standards on subsequent surveys. Moreover, interviewers ought to understand that some percentage of their calls will be checked—when checking is possible—not as the result of suspicion of her work, but as a *routine of sound research practice*. Another frequently overlooked device for improving the quality of interviewing is the sending to all interviewers who worked on a survey some capitulation of the findings. This inexpensive attention makes the interviewer feel that her work was a part of something important and that it really came to some result.

PERMANENT STAFFS OF MOBILE INTERVIEWERS

One recommendation for improving the quality of field work on consumer surveys is the building up of a small staff of highly competent interviewers, paid on a regular monthly salary, who can go from place to place in the course of a survey to complete the consumer interviewing. This plan, urged by Dr.

D. B. Lucas, has actually been used by Dr. Rensis Likert, chief of Division of Program Surveys, Department of Agriculture. There is no doubt that such a staff of permanent interviewers could turn in better work than part-time and inadequately supervised local interviewers. Its high cost, however, makes it seem unlikely of wide adoption.

OBSERVATION OF CONSUMER BEHAVIOR

All consumer surveys based on statements by consumers about their own actions or thoughts are open to some suspicion. Can the consumer conquer his tendency toward rationalizing his own behavior, toward putting his best foot forward? Can the consumer tell the truth about himself, or herself? On many questions of objective fact, the consumer can and does to hundreds of interviewers every day. But on some questions of marketing information, a much straighter story can be pieced together by observing the consumer.

How does the customer shop when she goes to a food store? Does the window display that attracts attention attract the people who stop to look at it into the store to look further, or to buy? Do most purchasers of magazines from newsstands know what they want to buy when they come to the point of purchase, or do they leaf through several magazines and make a selection then? How many residences in a given city or area need paint? Is the motorist influenced by the sales talk of the attendant to buy additional items at a filling station?

These and hundreds of similar questions can be answered better by observation than by interrogation. The market research worker can station observers in filling stations, in food stands, or near newsstands to watch and record what customers do—and that is a much more important record than a statement by consumers of what they think they do.

This principle can be carried further. Several years ago the Commercial Research Department of the Curtis Publishing Company devised an ingenious adaption of the idea in an experimental consumer shoe survey. Conventional approaches to the question of what brands of shoes women buy had yielded an undue proportion of higher priced and well-advertised shoes and a far too low proportion of bargain-counter shoes. So interviewers called on a small but well-rounded sample of housewives, first asking them what brands of shoes they had purchased, and then actually taking an inventory of the shoes owned by each respondent. It was apparent that the observation method yielded a much more accurate story than the interview method.

These—the mail questionnaire, the consumer panel, the personal interview, and the observation method—are the major ways in which the research worker can get direct information about the behavior of the consumer. In

most instances, the method will require asking the consumer to report at least some part of the information needed. Therefore, the way in which questions are asked must be given serious attention.

BUILDING THE QUESTIONNAIRE

The first requisite of a good questionnaire for a consumer survey is that it be possible to answer. At first glance, this sounds like a truism—yet many questionnaires are so out of tune with the *thinking of consumers that consumers really do not know how to answer the questions. To avoid this difficulty, there are several pretty well-established principles that should be followed.*

Questions, so far as possible, should be specific and objective. Complicated and speculative questions are likely to be *misunderstood and to bring back rationalization rather than fact.* A conditional question, "What would you do if . . ." is *not only difficult for the consumer to answer, but equally difficult for the analyst to interpret.* Questions should ask about simple facts, recent actions, and should avoid generalization. Such a question as "What do you usually eat for dinner?" can be answered succinctly by few consumers. But the question "What did you have for dinner yesterday?" can usually be answered.

Many times the object of a consumer survey is to learn whether a new product or an improved product will be accepted by the consumer. Yet little can be learned of what the research man really wants to know by asking consumers questions merely about the new or the improved product. If, however, the interviewer is able to ask for a comparison between two products, which can be used under reasonably comparable conditions, she may be able to get information that will reflect genuine consumer preference. If the objective of a consumer survey is to learn whether consumers prefer a new blend of, say, coffee, the interviewer should take out samples of the new blend and samples of various popular brands with which the new blend will compete, but in every case with the brand identification removed. Then, when a consumer reports using the popular Brand X, the interviewer can ask her to try out and compare two coffees, one, the blend under consideration, and the other, a disguised sample of Brand X.

Questionnaires should be short. It is rarely necessary to ask a great many questions to adduce the information needed. Interviewers have frequently been asked to obtain answers to questionnaires running to the length of 12 or 15 pages, and requiring well over an hour to answer. The interest a consumer will take in an interview is about in direct proportion to his or her interest in the product or the subject of the survey. And unless only people

with a genuine interest in the subject of these jumbo questionnaires are interviewed, results may be more misleading than helpful.

The best way of determining the length of a questionnaire is for a man or woman who develops it to do enough experimental interviewing on it to be sure that it is workable. In fact, the busy market research executive could profitably spare a portion of his time to do a day's interviewing on the questionnaires pertaining to every important survey that he directs. If this were standard practice, there would be far fewer unwieldy and vague questionnaires and far fewer discouraged interviewers.

Experimental interviewing for all but the simplest of questionnaires is virtually a necessity of questionnaire building and is usually a routine part of procedure under experienced market research technicians. When a problem has been formulated, a preliminary questionnaire is usually whipped into shape in conference and then sent out to be the object of experimental interviewing. When the faults and inadequacies are discovered by this experimental interviewing, the questionnaire is redrafted and again submitted to a field test. This process is repeated as many times as necessary to develop a questionnaire that can be worked smoothly and will adduce the information needed.

It may be said that this method of trial and error in questionnaire building is both time consuming and expensive. It certainly is, but it is less time consuming and less expensive than rushing to the field with a questionnaire that does not answer the needs of the particular survey.

In the early stages of the development of the questionnaire, it is sound practice to plan the tabulations to be used when the returns are all in. It is easy to think of dozens of questions it would be pleasant to have answers to, and which might conceivably bear on the problem at issue. But unless these questions can become a logical part of the tabulation plan they will be useless. And when the specific tabulations to be run are planned in advance, it soon becomes apparent what questions are superfluous—and sometimes what question not originally included will be needed.

For surveys including several thousand interviews, many research technicians recommend precoding for machine tabulation. Precoding requires putting on the questionnaire form a symbol for each answer which may be given to each question, so that punch-card operators can punch the tabulating cards for machine tabulation as soon as the interviews come in. There is no doubt that this practice saves a great deal of time in tabulating questionnaires. And it must be admitted that questionnaires must be edited and coded when they come in, if they are not precoded. Like answers stated differently should be thrown into the same categories. And there must usually be an "All Other" category into which all the cats and dogs of answers have to be thrown.

But there is one serious disadvantage to precoding. On surveys on which the

24 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

number of possible answers can not be known in advance, precoding tends to limit the scope of the information adduced by the survey. If a survey is made to determine the age and brand of washing machines owned by people in certain-sized cities (where possibly retail distribution may have been different for a particular brand), it is true that a dozen or fifteen brands will lead. But in various such surveys, it has been found that consumers will be using several hundred different brands of washers. Now, precoding would have obscured this condition and kept out of attention the apparently large number of retail outlets that have sold little-known or private brands of washers. Only when the range of answers can be forecast with considerable accuracy is it safe to resort to precoding, and in that case it is highly desirable.

There is one other objection to precoding which needs to be mentioned. Precoding is exceedingly difficult to explain to interviewers. Typically, interviewers know little or nothing about *machine tabulation* or about *tabulation* of any kind. And getting them to throw into the "All Other" classification all instances not listed by specific name is difficult. In some surveys, although precoding may be justified on all other counts, the difficulty of explaining it to interviewers may be a definite deterrent.

LAYING OUT THE SAMPLE

The skill with which the *research technician* lays out his sample—that is, designates how a true cross section of the group he intends to study can be located by interviewers—is another of the basically important aspects of market research.

The first question that comes up is "How many people must be interviewed?" There is no general answer to this question. It depends on the nature of the information sought. If the researcher wants to learn whether it hurts to sit on a tack, he does not have to slip a tack on the chairs of many people to find out—and he can find out entirely by observation! On the other hand, if he wants to find out what brands of coffee are consumed by the people of the United States, and in what proportion, he must—considering the hundreds of private brands and regional brands and national brands—sample every area in the country; he must sample all income levels; he must sample all racial groups. And he must sample them in sufficient numbers to avoid any unusual or atypical conditions in any of these areas or among any of these groups. It would require a considerable number of thousands of interviews to produce reliable results. But if he merely wants to learn the proportionate share of the market held by several leading national brands, he can do so with a much smaller number of interviews, possibly 2,500 to 3,000.

There have been a number of so-called short cuts to figuring out the point

at which a sample stabilizes, that is, when enough interviews have been made truly to reflect the entire group under study. Sometimes a relatively small number of interviews will be tabulated—say 500. Then, if the addition of another hundred interviews does not change the profile of the tabulations, the sample is regarded as stabilized. Sometimes running tabulations of certain questions in a questionnaire are kept, and charted, and when the graph stops erratic behavior and levels off, the sample is regarded as stabilized. The difficulty with these short cuts is that they work on some kinds of problems and not on others. Probably the market research technician who knows enough of sampling techniques to apply these short cuts without risk knows enough to get along without them.

In general, it may be said that questions bearing on physiological response require a relatively small number of interviews. If the research worker wants to find out whether a particular detergent has an odor that will prove to be offensive to potential purchasers, he can find out by having only a few dozen housewives smell the detergent. Or if he wants to find out if the layout of an advertisement emphasizes certain sales points, he can find out by putting the advertisement before only a few dozen literate people.

But when he wants to obtain information that will allow him to generalize on the behavior or the attitudes of a large group of people, he will have to resort to a sample large enough to comprise all types included in this group and to comprise them in about the right proportion.

It is probably fair to say that a knowledge of sampling techniques is just now emerging and that the knowledge of sampling has not yet developed to the point where widely applicable generalizations as to necessary size of samples can be made.

TYPES OF SAMPLING

There are two general approaches to the problem of laying out the sample. The first is called purposive, selective, or ratio sampling. The second is called *area* sampling.

The first type (purposive, selective, or ratio—any one of these three terms may be used) is the older method of determining what individuals should compose a sample. If, in a given locality, 85 per cent of the population is white and 15 per cent is colored, then the two groups should be represented in the same proportion in the sample. And if 50 per cent of the population is affiliated with one or another of the Protestant churches, 25 per cent with the Catholic Church, and 25 per cent with no religious body, then the same proportions should be maintained in the sample.

The chief difficulty with such a procedure is that it is difficult to control the sample on more than two, or at the most three, breakdowns. If this technique

is to be applied rigorously it becomes cumbersome. On the other hand, if the group to be studied can be described by a simple set of ratios, then this method will usually work satisfactorily. If, for instance, a study is to be made of brand preference among the various gasolines, it might be adequate to control the sample by maintaining the same ratio of ownership of various brands of automobiles in the sample as the total group shows.

The newer technique, area sampling, is undoubtedly much more accurate. It is also much more costly! Area sampling has been developed largely by the United States Bureau of the Census as a method of obtaining needed information concerning the population of the country between the periods of the decennial censuses. Under this plan, a random selection of a small part of the land area of the country is made and all the people in each of the resulting areas are questioned. At present, 2.8 per cent of the land area of the United States—made up of many parts of counties scattered all over the 48 states—is being used. But it costs approximately three million dollars to put a questionnaire through this process—a sum beyond most commercial research budgets.

A word must be said about the term "random." This term has primarily a statistical connotation. A random selection is a selection in which any individual in the total group has just as good a chance of being selected as any other. When a coin is tossed in the air, "heads" has as good a chance of coming up as does "tails." And when Selective Service "drew" numbers to determine the order of calling, each capsule containing a number had as good a chance of being drawn first as any other. When land areas are selected at random, they are numbered, and the numbers are selected by a method that assures as great a chance of selection for one area as for another.

Under the area method of sampling, the areas that happen to fall in, say, the state of Oregon would not be an adequate basis for sampling that state. It might happen that none of the areas fell in the city of Portland, which has almost 50 per cent of the population of the state. If the state of Oregon were to be sampled on the area basis, it would be necessary to approach it the same way that sampling of the country as a whole is approached.

The sampling work of the Census Bureau shows great promise and may result in notable improvements in the availability of census information and in considerable advances in market research technique. At the present time this method, in its pure form, is too costly for most commercial research work.

The better plan is a compromise method. If a single city is to be sampled, the research worker can take a map of that city and grid the city completely by means of horizontal and vertical lines at equal distances, dividing the city into possibly 40 approximately equal areas. He can—if the survey is a consumer survey—exclude purely retail, commercial, and industrial areas. Then, if he has 25 areas left, he can relentlessly require 1/25 of the interviews he

plans to obtain to come from each area. When he has his interviews in, he can test his work by comparing the ages of respondents or the occupations of the major wage earners of each family with similar information from the census reports.

This method of sampling is reasonably simple to apply and yields remarkably adequate samples if several hundred interviews may be obtained from the city sampled. It must be admitted, however, that experience in the field of market research is more valuable in laying out a sample than in any other aspect of the field.

ANALYZING THE RESULTS OF THE SURVEY

Three kinds of results may come out of market research projects: (1) information, (2) interpretation of that information, (3) recommendations. When the returned questionnaires are tabulated, there come to the analyst a number of pages of tabulations—usually complicated in appearance and requiring considerable study. Clearly, this is not the type of material to put before management in its crude form.

Probably the analysis of tabulations is more of an art than a science, and experience will teach more about how far it can be carried than precept will. But the following suggestions may be useful.

It will be remembered that the research man went to considerable trouble at the outset to formulate a concise statement of the problem that necessitated the survey. He should put that statement before him and study the tabulations to see every fact which can be drawn out which will bear on that problem. If he has participated in all parts of the survey, these facts will be more significant to him than to anyone else and he will be in a better position to formulate whatever is significant about each of the facts which could be adduced from the tabulations. In most cases, it is not the facts that are of paramount importance—it is what the facts mean. If he wants to be systematic—and it pays to be systematic at this stage of the research project—he will make a formal list of the facts he has learned, and under each fact will write down what significance or interpretation he can see. From these facts, with their interpretations, it should be relatively easy to make recommendations, if doing so falls within his province.

The report he submits must be easy to use and just as easy to refer to later. It is usually a good idea to include in the report a short statement of the research techniques followed, a copy of the questionnaire or questionnaires used, a statement of where the work was done, of when it was done, and by whom each part of the job was done. At the time a survey is freshly completed, this information often seems superfluous. But several years, or

even several months, later the report may be useless for comparative purposes without this information.

A short summary should highlight the facts found and the interpretations placed on them, with whatever recommendations are made. And it is usually good practice to append the pages of tabulations. It is true that few busy executives will take the time to reanalyze the tabulations. But they may be invaluable for later reference.

This discussion of market research can be summarized pointedly. There are fairly well-developed techniques for market research. But at no time can they take the place of judgment and ingenuity. The variety of problems that come to the desk of the market research technician is often bewildering. And, usually, different types of problems require different methods. The techniques are important. But judgment, careful planning, meticulous attention to details, and constant supervision are just as necessary. And if these are all exercised, market research workers can make an important contribution to the well-being of our entire population.

Chapter III

RESEARCH FOR PRODUCT DEVELOPMENT¹

PERCIVAL WHITE AND MATILDA WHITE

Market Research Company of America

MOST manufacturers find it necessary to improve old products or to bring out new ones. In the past, the usual procedure in developing a product has all too frequently been a direct route from idea to production without any attempt to predetermine acceptance by the public. This trial-and-error approach has resulted in many failures and tremendous cost to industry. Research methods now available and constantly being improved can, at a relatively small cost, avoid this type of waste by providing an orderly and scientific procedure for developing products that will have a definite market.

The principle of this method is that *consumer or user needs must be studied while the product is being originated and developed*. This means that the creative process is simultaneous with the testing, which proceeds on a scientifically planned basis, on a time schedule, and preferably on a small and intensive rather than on a far-flung scale.

Product testing among consumers is the counterpart of product designing in the laboratory. Designing and testing, redesigning and testing again and again—these are the steps in product development. The developing product zigzags from the laboratory to the test market and back again, through successive stages, until it finally emerges complete.

In general, the procedure for developing a product may be described as having six steps:

1. Choosing the product field, if an entirely new product is under consideration.

¹The material in this chapter is a syllabus of a forthcoming book on product development, to be written by the authors of the present chapter.

30 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

2. Analyzing potential demand for products in that general field.
3. Conceiving some ideas for a new product in that field.
4. Laboratory designing of product or products on the basis of these ideas, and making models or samples.
5. Testing user acceptance of these models or samples.
6. Making all changes and improvements necessary to ensure user acceptance in advance of actual preparation for production.

CHOOSING THE PRODUCT FIELD

It is of paramount importance at the outset that the manufacturer who wishes to launch a new product decide wisely upon his product field in general. The process is one of successive elimination, based upon facts, from all possible fields to a few fields, and from a few fields to a final choice of the product field best suited to his needs and opportunities. When a single field has thus been chosen, there follows the continued elimination of many possible products within that field until he settles upon the final choice of one particular product (or a group of products). This is done by collecting all relevant facts and then sitting down in executive meetings to weigh these facts.

The first step in this procedure is a general preliminary study of all the reasonable possibilities. At this early stage the manufacturer must decide whether to build and sell end products; to build and sell components; or to compromise, perhaps by designing and patenting an end product, giving free rights under the patent to someone else as long as he purchases the components.

The choice of field is affected by many factors, concerned with both the internal and the external aspects of the firm and the market. Although most of them are interrelated as cause and effect, they may in the main be considered separately. All factors should be weighed in making the decision, but the weight given to each factor cannot be stated arbitrarily or generally. This depends on the company itself and on its individual problems. A single factor may prove of such importance in a given case that it will automatically rule out products which are desirable from all other points of view.

Some of the principal factors to be considered are:

1. *Size of company, and capital available.*
2. *Manufacturing facilities (including machinery and plant facilities, labor supply, geographical location, and transportation facilities).* It is to be expected that a new product will call for some changes in manufacturing facilities, but it is desirable to accomplish the change with a minimum of

wrench and dislocation. The farther afield a new product is from that normally manufactured the more hazardous it is as a business venture, since the problems in every department of the business become strange and more troublesome.

3. *Past and present products and the reputation of the firm.* The new field can rarely be evaluated as a wholly distinct and separate entity. It must be considered in relation to the manufacturer's past and present products, their price levels, quality, need for service, fields of use. Past and present products have determined not only the facilities available, but also the reputation of the firm. It is, to be sure, profitable to make full use of an established reputation, when it is sufficiently good to be exploited in the new market. Although it is sometimes hard to convince the market of a company's ability to succeed in a field radically different from that in which its reputation is established, reputation is, nevertheless, a tangible factor with a definite dollar-and-cents value.

4. *Sales organization.* The question to be answered is: What changes will be called for in the sales organization? A shift from a consumer field to an industrial field, for instance, may demand additional technical knowledge on the part of the sales organization. A shift from an industrial field to a consumer field demands salesmen who can make contacts with an entirely new group among wholesalers and retailers. In changing from a specialized industrial field to standardized industrial goods, the company may find that its present salesmen are specialists, too high-priced for a less expensive field.

5. *Distribution channels.* Although it is usually possible to develop new outlets and many companies do not hesitate to do so, the use of existing sales outlets is frequently of great weight in the choice of a product field. One important reason why most companies confine new product development largely to rounding out their line is that this policy enables them to make efficient use of their present distribution organization.

The problems of the channels to be used are closely tied in with and qualify decisions on price, quality, and service.

6. *Competition.* Probably, complete information about the competition in the proposed fields is the most important single aspect of the preliminary facts needed to choose a product field. A study of the competition provides clues to the problems in that area and serves as a base with which to compare the various aspects of the company. The factors listed above must be evaluated not merely in themselves, but in their relation to the competition in a proposed field. A thorough study of the competition will not be limited to the competition on similar products alone, but will include that offered by other types of products. The manufacturer of flexible shafting, for example, faces competition not only from other manufacturers of this product,

but also from developments in pneumatic power, hydraulic power, electric motors, and gearing.

It is essential to know:

- What companies constitute the competition.
- Their volume of business.
- Their number and their location in relation to the size and geographical distribution of the market.
- The amount and extent of their distribution.
- Their method of operating and the channels of distribution they use in relation to the buying habits of the market.
- Their rate of growth.
- Their special advantages in the way of patents, etc.
- Their policy on advertising.

Their products, catalogues, advertising, credit ratings, etc., should all be analyzed as indices of the market.

It should be emphasized that every company has an individual problem. Rarely do all the above factors (and there are many other pertinent ones, such as raw materials, by-products, etc.) have equal importance in a given situation. The emphasis is merely on the need to base the decision on facts, on all the facts which are relevant and available.

A decision not based on facts, or based on incomplete facts, has often resulted in dismal failure:

There is the case of a new food product, for example, sweet potato flakes, "invented" by a producer whose experience had been limited to the sale of established products.

On a laboratory basis these sweet potato flakes were easily made on a small, experimental rotary drum dryer. They were delicious to the taste, and, carried away by enthusiasm, the producer invested many thousands of dollars in the project.

When production was started without further experimental work and with no special market testing among consumers, trouble began. The factory scale equipment did not work out. An atmospheric drum dryer had been built by the local machine shop without adequate technical experience and the flakes, instead of peeling off in one revolution, stuck to the surface of the drum. Re-designed and converted into a rotary vacuum dryer, the flakes came off satisfactorily and the product was launched without any type of consumer or retail store test or market study. The package was designed by the inventor of the food, and without tests for eye or shelf-appeal. Sales were slow.

A local advertising campaign did move some of the product before the next difficulty developed. Without a moisture-proof liner on the inside and outside of the carton, the hygroscopic qualities of the flakes drew moisture and created early spoilage. The product was withdrawn—loss \$50,000. Careful, thorough, and intelligent research could have secured all the pertinent facts at a reasonable cost.

ANALYZING POTENTIAL DEMAND

Once a general product field is selected, its market potentials must be studied. The manufacturer will find that a broad market study is necessary here, using materials from two main sources:

1. Bibliographical material from published, secondary sources, including statistics from government agencies, trade journals, trade association reports, reports of private organizations and foundations, etc.
2. Primary material gathered firsthand from potential users, dealers, competitive producers, and other useful sources.

An orderly method of analyzing potential demand begins with *secondary sources* and the facts available from them. This determines roughly the scope of the market. Beyond this, tentative hypotheses as to scope, limitations, and factors controlling the market may be established logically.

On this basis, some *preliminary interviews* with consumers, jobbers, dealers, etc., may be made. Sufficient informal interviews at the beginning of the survey will provide clues to the kind of respondents who should be included in the study, to the pertinence of the various market factors tentatively set up, and to other elements which may not have been considered. The best informal method is the intensive interview, without a definite outline of the information to be gained, in which the interviewer asks just enough questions to create an opportunity to listen. From this information a questionnaire may be built, tested, and applied in a broad market survey.

The *broad market study* usually estimates demand in the proposed field by studying:

1. Actual use of, and experience with, similar products or products for comparable usage.
2. Attitudes conducive to or tending to prevent further use of this type of product.

The user of a similar product can describe what kind of person he is, why and how frequently he buys products of this type, how and when he uses them, how much he uses. He thus provides clues to various conditions or situations which induce use of the product. Analysis of such replies indicates the size, scope, nature, and direction of the market.

Such a study is planned to fit each particular product field, but follows, in its basic outline, the general principles underlying all broad market studies. Sampling and analysis must be geared to measure all the important factors which control the market. In planning, it must be remembered that this market is either local, state-wide, national, or international; that it is urban

or rural, or both. Age and sex, for instance, limit the market for liquid stockings and razor blades; climate affects the sales of bathing suits and furnaces; seasonal factors have an effect on certain fields. Since readiness to purchase is not identical with ability to purchase, and since price is often the most important single limitation on the market of a product, the relation of living standards to the potential market must also be evaluated.

In order to deal adequately with all such factors, a major survey is usually required, involving several thousand interviews and several months of time. At the end of this work, however, the producer has mapped his course and knows both his possibilities and his risks.

CONCEIVING SOME PRODUCT IDEAS

Within this known market, the producer then has the problem of launching a product that will be competitively superior. Consumer suggestions of product ideas are of strategic importance at this stage. They tend to ensure user acceptance in advance. Careful research in the beginning will reduce the amount of experimental designing and testing necessary later.

Here, however, the research is usually small-scale, but intensive. Since the purpose of this study is to secure ideas, to stake out the territory within which the creative staff can concentrate its effort, a properly distributed cross section of the market is not necessary. Only enough interviews need be made to provide sufficient ideas. Three groups of persons may well be questioned: those who now use the same, or a similar product; those who have used them in the past, but have discontinued using them; and those who have never used them.

Most consumers are neither inventors nor designers. To draw creative suggestions from the average consumer, therefore, requires a highly skilled technique of questioning.

In searching for new ideas, the interviewer begins with a discussion of all the aspects of actual experience which might conceivably have a bearing on the problem at hand. If, for example, the aim is to design a new beer container for home use, the investigator questions housewives (or, in some instances, the man of the family) in beer-drinking households, touching on such points as:

- Proximity of retail source, and whether or not beer is delivered.*
- Habits of storing beer, and the problems involved.*
- Occasions of serving beer, and how it is served.*
- Etc.*

Such a picture defines the role that the beer container is expected to fill.

The interviewer then probes further, encouraging the respondent to con-

sider his previous experience with different types of beer containers and to point out any objections to, or *difficulties with*, them. People are usually better able to criticize than to create, but even from *negative comments* creative ideas may often be drawn.

Finally, the interviewer shifts the questioning from past experience to the ideal container. (In the case of some entirely new product ideas, there is, of course, no past experience to be discussed, so that the interview can only be started at this point. But "entirely new" products are rare.) Here the trained investigator follows such rules as these:

Allow the respondent time for thought and considered judgment.

When he (or she) begins to run out of ideas, prompt the respondent, not with leading questions, but with probes that will encourage him to talk further.

Waste no time with those who lack imagination and have nothing helpful to offer.

Recognize and encourage the intelligent and creative respondent, and probe for ideas. With such respondents a trained interviewer spends more time, and asks permission to call again.

From this study of consumer habits and suggestions, certain rough ideas will emerge which should contribute not only to the design of the new product, but also to the whole production and marketing program. Such ideas will, as a later step, be submitted to consumers for test.

For products difficult to present for testing, this collection of consumer ideas is of paramount importance. It is relatively simple to produce samples of a whole series of different cereals or face creams or women's slips, and then to test each of these in actual use and so to select the preferred types. Such a trial-and-error process is not simple when it comes to washing machines or road rollers.

In general, major attention must, therefore, be given to the *advance* collection of consumer suggestions in developing products in the following fields:

1. Long-lived items, such as consumers' durable goods and most major capital goods and equipment.
2. Other items which are costly or complicated to design and to prepare samples.
3. Items which cannot be *freely experimented with*, such as those medical products (laxatives, contraceptives, etc.) which, though *controlled* by physicians' standards, must depend on consumer acceptance for their sales.
4. Services.

DESIGNING THE PRODUCT OR PRODUCTS, MAKING MODELS OR SAMPLES

After the study of user needs and ideas, the experience and suggestions of the consumer are pooled with the technical skill and inventive genius at the disposal of the manufacturer, in order to make the product that will sell at the greatest profit.

Consumer ideas are for the most part negative, and the inventor, designer, or engineer must turn them into positive and dynamic form. A study of consumer experiences with desk pens revealed, for instance, that consumers had five main objections to desk pens.

1. Desk pens are too fancy in design.
2. Everybody else uses the pen, because of its tempting accessibility.
3. It is another item to clutter the already crowded desk top.
4. The ink tends to dry out.
5. Desk pens often leak.

Objections such as these provide the company's creative department with, as it were, a blueprint in reverse of a desk pen that would be satisfactory to consumers. In other words, a design that overcomes these objections is presumably assured of popular favor.

In this instance the inventor, by various ingenious devices, accomplished just that. Thus the pen, now patented, was planned aesthetically and mechanically to overcome the objections of the consumer to similar products now on the market.

While the ideas are still in the stage of rough specifications, it is often desirable to make a patent search. The value here is threefold:

1. Issued patents may suggest ideas which, though never developed, have latent promise.
2. Ideas which appear to be new are often found to have been developed long ago.
3. Strong current patents may come to light which make it unwise to proceed for fear of infringement.

After the creative department has developed two or three designs, it is usually desirable to subject these to a final set of tests. Which of the two or three designs is most acceptable? What improvements are indicated by test performance? For this purpose, actual samples may best be used. Where this is not feasible, tests may be made of working models, photographs, diagrams, or sets of specifications.

TESTING USER ACCEPTANCE

Products, once they are developed by this empirical process, may be tested in the laboratory, in actual use, and/or at the point of sale.

The uses of *laboratory testing* to a manufacturer include the establishment of a factory standard, knowledge of the claims that can be made for his product, directions for the use of the product, etc. In the laboratory the manufacturer can weed out those designs which are not practical, thus getting down to a few styles or models or formulas for his market tests. Laboratory tests do not, however, take the place of performance tests among users. The area that a gallon of paint will cover in actual use depends not alone on the quality of the paint; it depends also on the brush used, the skill of the painter, the condition of the surface, the relation of the new to the old color, etc.

Tests in actual use should always be made in addition to, or in conjunction with, laboratory tests. They allow the product to be perfected, and its acceptance estimated, before it is even marketed at all.

A use test should be set up with the same scientific accuracy as a laboratory test. It should meet three basic criteria of laboratory accuracy:

1. It should deal with facts, not with guesses. It should not ask questions about an unknown and hypothetical future situation, nor should it discuss, by way of comparison, existing products whose actual use by the respondent is already half forgotten.
2. It should study conditions approximating those of normal, everyday use. To set a person up as an "expert," for instance, and to ask him to make repeated tests of various products, is likely to give an artificial result.
3. It should isolate each of the many factors involved, so that each one may be evaluated individually. It is practically impossible to study price, package, brand, and inherent quality all at one fell swoop.

The procedure of setting up such scientific conditions for product testing may be outlined only in a general way because the procedure must be modified, sometimes drastically, to fit each product. A flexible shaft or flatiron cannot be tested by the same technique as a pudding mix. Although the method is equally adaptable to well-established products, examples here will also be confined to the testing of new products because of their pre-eminence today.

Before the actual test can be started, several basic decisions must be made:

First: Should the test consist of a single call or must the product be tested over a period of time? It used to be common practice for investigators to ask

housewives merely to take a taste first of one brand of coffee and then of another. On this basis, she was expected to make a choice. Yet, this procedure set the respondent up as a sort of judge or expert and so was liable to result in fallacious answers. Still more important, it created conditions of use entirely unlike normal conditions.

To be sure, there are certain tests which may be made at a single call. If, for example, the choice is between a red tooth-paste tube as compared with a blue one, or one design of milk bottle as compared with another, a single call may be sufficient. As a general rule one-shot judgments are adequate only:

- a. When the test is one of appearance.
- b. When past experience is without any doubt adequate to allow critical evaluation of the new product. In the case of a flat milk bottle, for instance, every housewife knows whether or not she has difficulty finding adequate storage space for milk. Or in the case of many industrial products, the potential user may have sufficient experience to base judgments on blueprints or working models.
- c. When the product is expected to sell on impulse, and repeat sales are unimportant.

Single-call tests should always be regarded with suspicion since in many instances unforeseen advantages or disadvantages will appear when the product is actually used over a period of time. The apparently desirable flat milk bottle may, for example, prove to spill the milk in pouring. Consequently, in most cases it is desirable to test a product over a period of time. A new flour or a new flatiron can certainly be evaluated only after it has been put to actual use.

The second decision to be made is: Who should be tested? To test a product on a cross section of people means that many people may be included who will never actually buy such a product at all. Therefore, products are usually tested on people who already use the same or a similar product. The assumption is that all future converts to the product will be similar to present users in their reactions.

If a milk bottle is to be tested, its test market may roughly be assumed to be everybody, since most households buy milk at some time or other. If the product to be tested is a diaper or a new type of tobacco, the test users should be, respectively, mothers with babies or men who smoke pipes.

One great advantage of selecting the test users in this way is the resultant homogeneity of the sample. Because of this homogeneity, small samples of users will often yield highly indicative results. Samples ranging from 50 to 200 users are often adequate for any single comparative point.

A third decision which must be made in advance is: What products should

be used as controls in the test? Wherever possible, the test product should be compared with similar product. This control product will in most cases be one which is already established and is expected to be the closest competitor to the test product. Thus, a new shirt will be compared with a well-known brand of shirt, a new toaster with the commonly used design, etc.

In this way it is possible to tell not merely whether or not the people like the new shirt, but also whether they like it better or not as well as a shirt for which the market is already a known quantity.

In some cases a single new product may have to be tested against each of a series of established products because the new product will have to face several different competitive situations. Because of geographical differences in syrup usage, for instance, a new syrup might have to be tested against maple syrup in New England, against corn syrup in the Middle West, and against molasses in the South. Or, to take another example, a new tooth paste might have to be tested against Pepsodent among Pepsodent users, against Kolynos among Kolynos users, and so on, since a new tooth paste would be competing not with a single product but with an entire group of established brands.

A fourth decision to be made in advance is: How can the test problem be broken down into its various aspects? Every product has a series of aspects, any one of which may in itself be sufficient to constitute the success or failure of the product. In general, the aspects of a consumer product may include, for instance:

1. The inherent quality of the product, such as its flavor, effectiveness, etc.
 2. The type of package.
 3. The brand name.
 4. The price.
- Etc.

In so far as possible, each one of these factors should be tested by itself while all other factors are held constant. A test would be inconclusive if it compared a new soap powder with *larger* granules in an *oblong* container, on the one hand, or with an established soap powder with *smaller* granules in a *square* container, on the other hand.

The resultant preference might be due *either* to the container *or* to the size of the granule, or to both. In this instance, two separate tests should be made:

1. One test comparing the packages by keeping the granule size the same.
2. A second test comparing the size of the granules by using identical packages.

40 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

In testing a new soluble coffee, to take another instance, some four separate tests might well be made, since the new product might differ from established competitors in at least four respects. One test would compare the mild flavor of the new product with the stronger flavor of one established competitor. A second test would measure consumer reaction to the addition of carbohydrates in the new product. A third test would package the same product in two different types of containers in order to gauge package preference. A fourth test would measure the relative resistance to the new and unpromoted brand by stating on the packages the actual brands of the products contained: one new brand name and one established name.

The specific steps involved in one type of use test² may be illustrated with a new soluble coffee product. The problem to be studied, let us say, is whether or not carbohydrates should be added to the new product. The test products, labeled so as to make this content difference clear, consist of packages of the new coffee and an established soluble coffee without carbohydrates.

Investigators call upon a group of families several times at stated intervals, providing them with the necessary test supplies, observing the order in which they use the various packages and measuring the rate at which they consume the new product as compared with the old. At each contact, investigators offer the housewife a choice between the two products in order to estimate preference. They also question her about the reasons for the preference, and from this, possibilities for product improvement may be deduced. At the end of the test period they allow her choices between several pairs of alternative gift offers so devised as to show roughly how much more she would be willing to pay for one product than for the other. They also ask various questions about the preferences of the different family members, the uses to which the product was put, etc.

In this way, the test approximates the normal situation which the new product will have to face: its initial chances at the point of sale, its success in actual use, its chances of repeat sales. Its rate of consumption may be measured against that of its competitor. Preferences between the two products are based not on guesswork, but on experience.

This procedure, as outlined, is best adapted to foods and drugs and other quick-consumption items. With modifications, however, its principles may be adapted so that most consumers' goods and some industrial products may be pretested in a situation simulating that of normal use. A woman's slip may be studied by providing women with two slips, one new and one old, and then offering a choice between the two after a period of wear. A new orange juicer may be planted in a drugstore beside the old one, and the

² First developed by Dr. Raymond Franzen.

relative merits of the two discussed later with the fountain clerk. A new magazine may be distributed for a trial period, and its readership then be compared with that of magazines ordinarily in the house. And so on.

It is usually unwise to launch a new product unless the evidence is clearly in its favor; a fifty-fifty preference between new and old means that the new product could be expected to equal its competitor in sales, but only after it had received equal promotional backing. A new product usually has difficulty equaling an established product short of an extended period of time.

Nevertheless, there are instances where the fifty-fifty preference may be neither final nor conclusive. If the preference for one of the two products has increased from 10 per cent at the first call, to 30 per cent at the second call, to 50 per cent at the third call, the assumption would be that its success might increase still further with a longer period of use. Such a product might be expected to make a competitive place for itself if given a long and intensive promotional start. The best test procedure, in this case, would be to prolong the period of the test, providing the families with larger supplies, and ultimately going back for a fourth call.

After the use tests have been conducted, and the product adapted to use needs, sales tests have great value in solving point-of-sale problems. They high-light in advance any retailer problems or resistances. They allow experiments with display, promotion, price differentials, etc.

The sales test, however, valuable as it is, should be used in addition to, not as a substitute for, the consumer test in actual use. This is especially true since a sales test has certain disadvantages:

1. Since the test must be fairly broad in scope if it is to have meaning, competitors have an opportunity to copy the best features of the product, or to build up sales arguments against it.
2. It is difficult to follow the transaction once the product is bought. In order to determine satisfaction in performance, chances of repeat sales need to be evaluated, and this information cannot be secured from a simple sales test. For these facts, a second process would be necessary.
3. It is difficult to give proper weight to the various factors affecting sale, such as weather, sales in other stores, too much or too little promotion, because of the complex of factors involved. Under conditions of short supply, it is practically impossible to interpret the findings.
4. The sales test takes considerable time and money.
5. A considerable quantity of units must be made up.
6. Some products are completely unadapted to it.

Many products, before they are finally launched, are subjected to all three of these types of test.

PREPARATION FOR PRODUCTION AND MARKETING

Thus, the process of designing and testing, redesigning and testing again, continues until a final design is determined. In an absolute sense no design is final, but it may be considered final for a particular period of time. At a given point, however, the design must be considered final, not because it has achieved perfection but rather because executive judgment says: "We can't afford to experiment any further; we must get into production and begin making profit. This design is known to be marketable, so we'll use it."

Preparation for production and marketing is a problem with many ramifications. It includes making all the necessary changes in facilities, personnel, etc., which were discovered earlier in the procedure. It involves such matters as:

1. *Test runs*, which are necessary because the production problems are not always apparent in advance. Enough units are made to determine how the manufacture of the product works in a mass-production setup.
2. *Packaging, trade name, and other merchandising problems*, many of which may already have been solved in the use tests.
3. Adequate *distribution*, which should be related to the coverage of the advertising, look forward to problems of service, replacement parts, etc.
4. *Advertising*. The budget should be carefully planned to permit adequate advertising and promotion of the product before it is "out of the red." Many of the advertising appeals and promotion angles will have been suggested in the use and other tests. The dealer should be sold on the new product, not only through the results of the use tests but by means of display materials, etc. It may be desirable to distribute samples, offer special premiums, or trial offers, in order to break down the initial resistance to a new product.

This method is the modern approach to product development. It conceives of the product in relation to its user as well as to its producer. It avoids the ancient error of going into production first, and then later suffering heavy losses which result from foisting off an oversupply on an unreceptive market. It utilizes the "laboratory of normal use," as well as the tools of physical science. Thus it applies all the principles of efficiency to the designing of the product, as well as to its marketing.

The strongest evidence that this method is sound is that it is being more and more widely used. The manufacturers who have employed it find that it works.

Chapter IV

MAKING A TECHNICAL INDUSTRIAL SURVEY

CHESTER E. HARING

Committee for Economic Development

SO-CALLED "industrial" survey research is similar in its problems to "consumer" survey research. On the other hand, industrial research does have several marked points of departure from the consumer survey.

Industrial research is just what the name implies: research among industrial firms. Of course the nature of the problem or the group to be surveyed may vary tremendously.

THE KINDS OF PROBLEMS

Just as in the consumer field, industrial research studies markets, the nature of buying habits, readership of publications and advertising, product research, and public relations.

Usually, of course, it is the manufacturer of industrial products and equipment who undertakes the study. Sometimes his agent for the study may be a publication, an advertising agency, or a private contractor specializing in that kind of work.

1. *Market research.* The maker of industrial equipment may need to know the size of his market, its location, the degree of its concentration in different areas, and its nature (i.e., whether breweries, bakeries, or some combination). In order to determine successfully that his product will be a success, he must determine the relative "brand" standing of the item he contemplates. He must be able to make a prediction as to the probable size of this market in later years.

In other words, his problem is similar to that faced by the maker of consumers' goods, except that he is working with a different group. Just as the consumers' goods manufacturer knows that his market shows concentrated

areas, so the industrial manufacturer knows in advance that a majority of industry is concentrated in a relatively small number of counties throughout the nation.

2. *The nature of buying habits.* Related to this question of markets is the issue of buying habits. The industrial researcher has one quite different problem here from the consumer researcher. The latter, in studying a family, generally has to locate his prime buyer from among 2, or at most 4 or 5 people. The research man studying industrial markets has considerably more trouble. It may be the individual in charge of plant purchases who is responsible, or it may not. The man may be called *purchasing agent*, or he may not be. There are literally hundreds of possibilities of the responsible person in large industrial firms. All of which means that the industrial research man must first make sure that he is interviewing the correct individual.

Once he has located that person, then there are many questions he may want to ask. He may ask about the method the buyer uses in selecting this particular item; he may ask about the nature and frequency of repetitive buying; in fact he may ask just about the same series of questions that the consumer surveyor does in studying buying habits.

3. *Readership.* Dr. Freiberg, in his chapter on copy testing, describes some refined methods of measuring readership of advertising. In the field of industrial advertising, we, too, have attempted copy testing, but nothing so elaborate as that discussed by Freiberg. In fact one might say that copy testing in industrial fields is in its infancy.

On the other hand, we have developed sound and tried methods of measuring editorial reading of industrial publications. These are basically the same procedures used by the consumer publications, and described in Ludeke's chapter.

4. *Product research.* Industrial research may be concerned with testing the product to determine what potential users want it to be like. For example, one large manufacturer of industrial machines is now quizzing users of his type of equipment about exactly what they want and need in postwar years.

5. *Public relations.* Again, industrial research may be concerned with measuring reactions of industrial concerns to various companies manufacturing needed equipment. While this type of activity has, again, not gone so far as the approaches suggested by Wulfbeck, nevertheless, this is attitude measurement in the nature of public relations research. The type of question may simply be "What manufacturer first comes to mind in this category?" but it is still public relations research.

6. *Government research.* While this book is properly not concerned with the "Washington-type" questionnaire, discussion of industrial research cannot be complete without a mention of this. Generally, this government ques-

tionnairing is concerned with the nature of the industry's business: its products, its employees, to whom it sells, and all the rest.

COMPARISON OF INDUSTRIAL AND CONSUMER RESEARCH

1. *Similarities.* Basically the same techniques are used in industrial research as those used in consumer study. The method of personal interviewing is probably used most frequently, but there are occasions when mail and telephone interviews are obtained. The questionnaire needs careful development and testing; plans for obtaining a "cross section" must be made; the sample must be adequate in size; the field work must be carefully conducted and controlled; the method of summarizing results is similar. In other words, the two approaches are fundamentally similar throughout all their steps.

2. *Differences.* At the same time there are some essential differences in the two approaches. There are differences in preparing the questionnaire, in setting up and obtaining the cross section, and in the field work.

The questionnaire, of course, is to be used with a highly specialized group of people, rather than with any cross section of consumers. This means that it must be phrased in terms that make sense to the respondent—terms which may often purposely be highly technical in nature.

In obtaining the cross section, there are also several differences from the consumer approach. Collecting the statistics in advance of setting up the cross section is one of these differences. Although the Census does help considerably here, the Census is probably even more out of date for industry than it is for the general population, since the conversion to war production. The industrial researcher, therefore, frequently consults some of the standard references, such as Moody's, in collecting up-to-date statistics. The difficulty here is that *none of the sources lists the small concerns; only the large ones are covered.* This means that probably 70 per cent or more of the statistics are up to date, but there is just no coverage of the other 30 per cent.

In setting up this cross section, one of the early problems confronting the researcher is the selection of the proper people to interview. If the product is heavy equipment, the title of the proper individual may be purchasing agent, chief engineer, plant superintendent, or any one of a hundred phrases. And even the definition of the same title in different plants varies. In the large plant the chief engineer is probably a quite responsible person, although the nature of his duties may vary widely from plant to plant. In a smaller plant his function may be nothing more than tending the fires and keeping the plant cleaned up. About the only way to overcome this problem is, in the case of personal interviewing, to have the investigator referred to the correct person; and in the case of a mail questionnaire, to phrase the letter so that the

questionnaire will be referred to the correct individual. Certainly no use of a purchased list, without a great deal of care, can be recommended.

So far we may have given the impression that it is always executives who are approached in the industrial survey. While this is generally true, on many occasions the lower classified personnel must be approached. Particularly in product research it is important to get to the people who make use of the equipment, in order to get their reactions for careful designing of the item.

Most consumer research, with the exception of the panel technique, does not require interviewing specific individuals. Rather, people are taken "as they come" within the restrictions of the cross section required. Industrial research, on the other hand, generally means that specific people must be interviewed.

The qualifications of the field interviewer, too, are somewhat different in the industrial survey. Not only must the investigator be of extremely high type, but for industrial work frequently must be a male. In addition, the training of the fieldworker is more detailed, for he must be so instructed that he can answer technical questions put to him by the respondent. The consumer interviewer generally needs to know little or nothing about the nature of the problem.

AN INDUSTRIAL PRODUCT RESEARCH STUDY

Charles F. Kettering once said, "A problem clearly understood is generally fairly simple." The proof of the truth of this statement was apparent when the Hooper-Holmes Bureau was retained by a manufacturer of construction equipment to conduct a study on the construction and operation of small power shovels.

The client was frank. His engineers were considering the design of a new power shovel of $3/8$ -yard and $1/2$ -yard capacity. The management found that it was encountering endless discussion on the details of design. They could not agree on the proper length of the crawler under certain operating conditions, they did not agree on whether the teeth of the ring gear should be on the inside or the outside circumference, etc. In short, the management found that its engineering department could not agree on whether it had wide or restricted latitude in dimensions on certain parts. At this point, the company employed our market research organization to obtain these specifications from the men who actually use small power shovels.

Hooper-Holmes inspectors, as their fieldmen are called, knew next to nothing about power shovels. Probably most of them knew that they always drew crowds ("sidewalk superintendents") in excavating in cities and caused detours on highways when the inspector was engaged in his daily work, but their knowledge went little further.

The truth of Mr. Kettering's statement was immediately seen. The problem was one of making entirely clear to our fieldmen the important facts about operation and construction of power shovels so that they might question the users without disclosing the meagerness of the inspector's actual knowledge and experience. Another important problem was to design a questionnaire that would develop the material desired and enable our nontechnical investigators to record it clearly and easily.

Since the actual results obtained from this survey can be of use to only a few persons, they will be eliminated from this discussion. On the other hand, the methods developed for instructing nontechnical investigators and the questions asked may be of help in industrial research. This discussion will be devoted entirely to these two phases of our study.

Because this industrial survey was so entirely foreign to the experience of the fieldmen, and would also be so to nearly all other research organizations, it seems instructive to quote at some length from our written instructions the portion we headed:

Background and Approach:

This survey should prove to be of considerable interest to all inspectors since it is in a field in which we only rarely are asked to work. And, since these respondents are not continually bothered by questionnaires like many housewives and retailers, it should be very easy. Of course all of you are familiar with power shovels which you see every day performing the work of large gangs of laborers, and we are sure that by the time you have completed your assignment you will certainly know more about power shovels than any of your acquaintances.

Our client is one of the important construction equipment manufacturers who is considering making small power shovels. Too frequently such equipment is designed by engineers with only hit and miss suggestions from salesmen and users. In this case, our client's engineers have agreed to let us make this careful study among the men who use their shovels. However, in this particular study, we are interested only in small power shovels—those of $\frac{3}{8}$ -yard and $\frac{1}{2}$ -yard capacity.

Although this study is national in scope, you will find your assignment is not large, and can easily be completed from the many names we are sending you. However, do not be disturbed if the listed individual is not available for he may be on the job hundreds or thousands of miles away. Our defense work has greatly upset the construction industry, and while ten years ago most construction contracts were awarded to concerns a few hundred miles away, today it is *not at all unusual* for a Pacific Coast concern to be building a powder plant in Maryland or Florida. For example, on the famous Pennsylvania Turnpike opened last year a Cleveland contractor obtained the contract to build one section, and he sub-contracted a tunnel to a Denver company.

But you need not confine your interviews to the individuals listed on the cards sent you. You may cover any responsible party holding a similar position on any

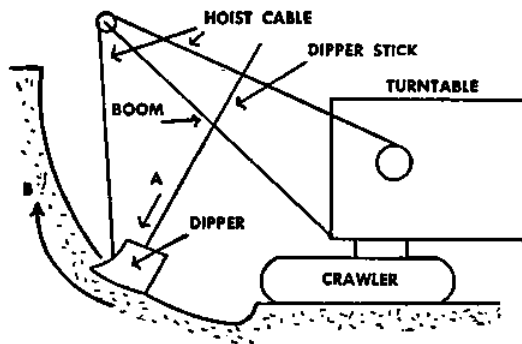
construction work in your territory, provided, of course, they are using *small* power shovels. In getting interviews with non-listed respondents, don't confine your efforts to your Branch City but get as many as possible from smaller towns nearby. General contractors are the large users. This is true because the *small machines*, which we are interested in and which we are discussing, are designed to be as versatile as possible and the owners of these machines very seldom restrict themselves to one individual class of work, but more or less are opportunists and will take any kind of job of excavating or material handling to which their machines may be adapted. They may be on a road job this week; they may be erecting steel this week and charging bins next week. Therefore, it is possible to classify them as general contractors. However, in many cases where we know that the use of the machine is along certain definite lines, this specific classification has been indicated on the card.

For the most part you will find the individual names given you are those of engineers on the construction job—rarely will they have office jobs. You will also find your quota probably includes a few operators—these are the skilled laborers who actually operate these *small* power shovels. We suggest that you try to get them during their lunch hour when you will find they will be only too willing to gossip with you. The engineer whom you have already interviewed will be glad to give you his name and ask him to talk with you.

In all cases make it clear to all respondents that our client is a builder of construction equipment who is thinking of diversifying this line and is, therefore, making this investigation of the $\frac{3}{8}$ -yard and $\frac{1}{2}$ -yard field to see if it warrants the development of a new and better machine. Also make it clear that you want to know the good points of all *small* shovels, but also the bad ones. Here is his chance to air all of his headaches and bellyaches at a time when they really will do some good. Undoubtedly the first question of many will be, "Who are you doing this for?" Your answer can only be that you do not know except that your New York headquarters said it was for one of the builders of construction equipment.

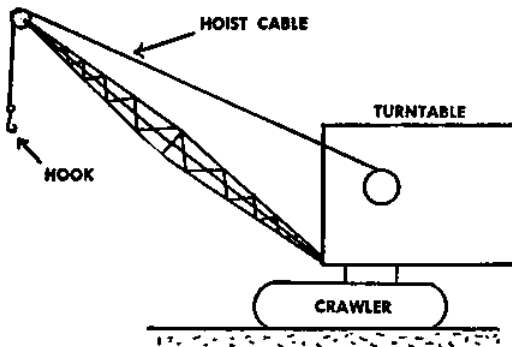
Construction of questionnaire. On questionnaires of this type and length, it seems advantageous to group together all questions relating to a given phase of the study, and to number separately each group of questions. Frequently the respondent will get a glimpse of the questionnaire being used. When he sees a total of 51 questions, as in this case, he is likely to object to giving the time. When each section was numbered separately, he could observe only 15. But the greatest advantage of dividing a questionnaire into sections is the aid it gives in making interviews. Instructions for this small power shovel study make this clear:

You will find this questionnaire is divided into several sections each covering one subject. We suggest that when you come to each section you say to the respondent something like this, say for Section C: "Now I'd like to ask a few questions about Turntables." This will help the respondent understand the meaning of each question and make him realize you are changing the subject. Frankly, we believe these questions are stated so clearly that your general knowledge would carry you through most



POWER SHOVEL

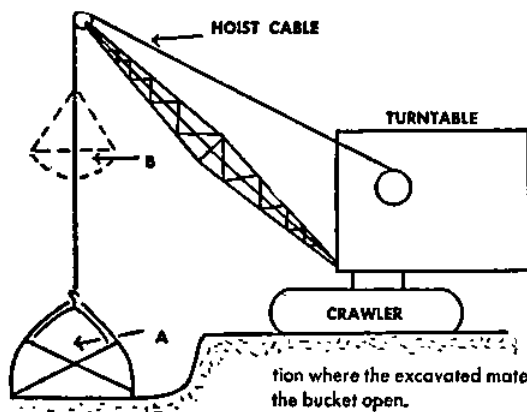
Digging is accomplished by pushing or thrusting the dipper into the material by a force (A) and pulling the dipper through the material (B) at the same time. Force (A) is the "crowd" and is applied to the dipper stick which is the handle or arm to which the dipper is attached. Force (B) is the "hoist" and is effected by pulling in on a cable which passes over the tip at the boom.



CRANE

Equipped with a longer boom than the shovel and of lighter construction, usually of latticed steel or of plain steel channels.

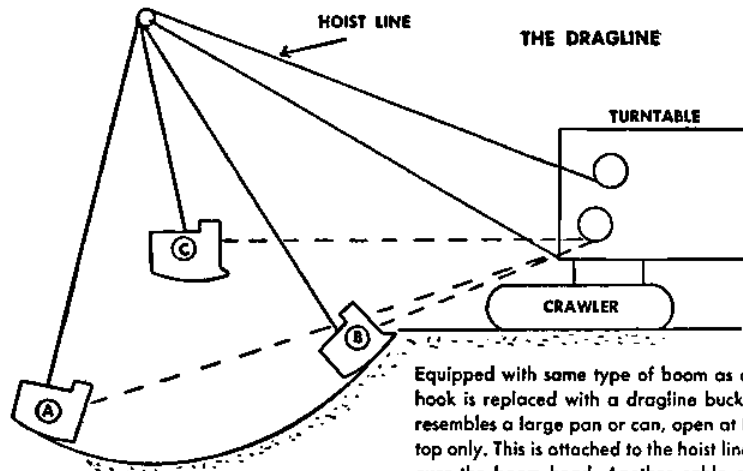
Hoist line runs over head of boom and is equipped with a hook on the end. Cranes are used for lifting service only in lifting pipe, tanks, boiler, structural steel and similar objects.



CLAMSHELL

Equipped with same type of boom as a Crane but the hook is replaced with a clamshell bucket, so called because of its resemblance to the two halves of a clamshell.

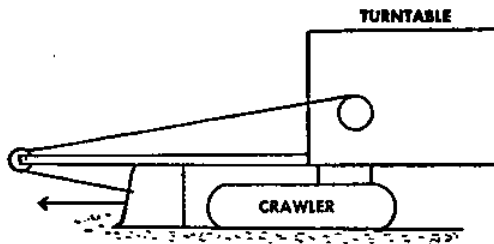
To dig, the bucket is dropped onto the material with the clamshells open, as shown in position (A). By pulling in on the hoist cable the two halves of the bucket are pulled together and are closed, becoming filled with the excavated material in so doing. The closed bucket is then raised, as shown in position (3) and swung to the desired position where the excavated material is dumped by releasing a cable and letting the bucket open.



Equipped with same type of boom as crane but hook is replaced with a dragline bucket, which resembles a large pan or can, open at front and top only. This is attached to the hoist line running over the boom head. Another cable called the drag-in cable runs directly from the bucket to another drum on the machine. (In sketch drag-in cable is shown by dotted lines.)

another drum on the machine. (In sketch drag-in cable is shown by dotted lines.)

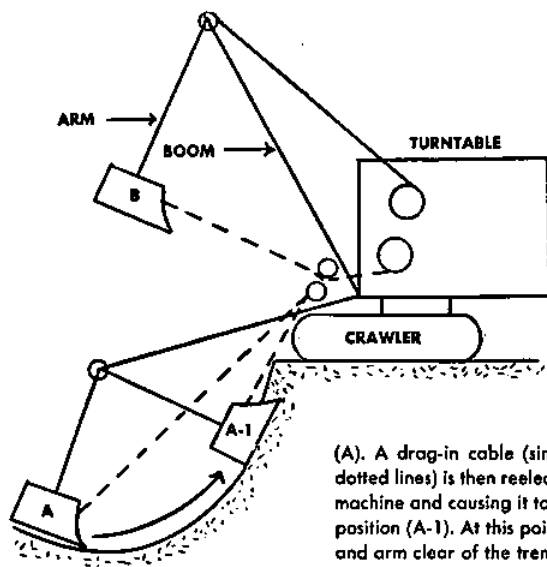
To dig, the bucket is dropped onto the material to be excavated as shown in position (A). The drag-in cable is then reeled in, pulling the bucket through the material from position (A) to position (B). In being pulled in in this manner the bucket becomes filled with excavated material. The full bucket is then raised into the air by means of the hoist line position (C) and swung to the desired dumping position where it is allowed to tip and discharge the material.



SKIMMER SCOOP

This unit has a relatively short boom which is in reality a track or rail, with a dipper along the bottom which rides back and forth on the boom.

To dig, the dipper is dropped to the lower end of the boom and the boom is then lowered to a position parallel to the ground. By means of a cable over the end of the boom, the dipper is pulled out or away from the machine, causing it to dig and fill with material. When the dipper reaches the outer end of the boom, the boom is raised into the air, and the dipper is swung to the desired dumping position where the material is then spilled out of the dipper.



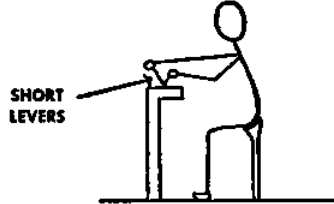
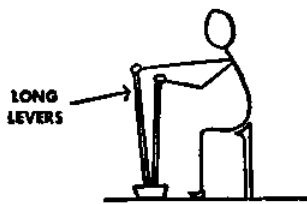
TRENCH DIGGER

As the name implies this unit is used primarily for digging trenches, and in some parts of the country (primarily Detroit) is being used for digging residence basements.

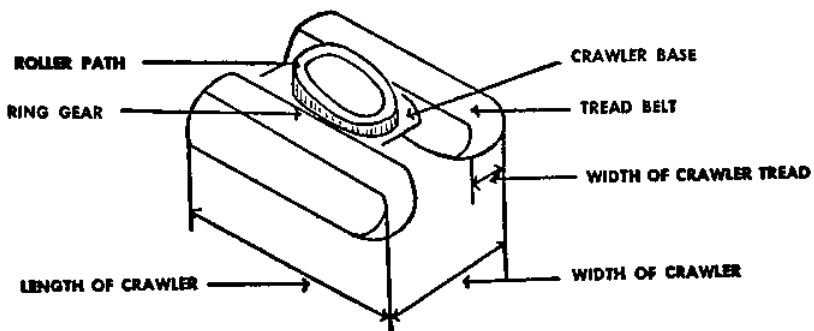
It is equipped with a stronger and shorter boom, similar to the power shovel boom, and has a steel, rigid arm hinged at the top end of the boom, with the dipper fastened to the lower end of the arm.

To dig, the arm is extended or swung out as far as it will go, and the boom and arm, in this position, are lowered into the trench, as shown in position

(A). A drag-in cable (similar to that of the dragline—shown by dotted lines) is then reeled in, thus pulling the dipper in toward the machine and causing it to dig and fill with material and to assume position (A-1). At this point the hoist line is used to raise the boom and arm clear of the trench to position (B). It is then swung to the desired dumping position and the material is spilled out of the bucket.



CRAWLER DETAILS

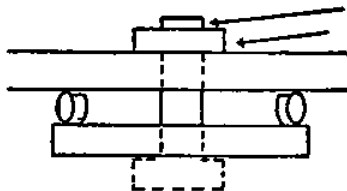


HOOK-PLAIN-NESTED ROLLER

The turntable revolves through 360° of swing and rests on top the crawler. To permit swinging the turntable it is equipped with rollers on the bottom which rest on a roller path on top the crawler base.

There are two methods of holding turntable and crawler together.

- (1) A center pin or pintle which is anchored in the crawler and extends up through the turntable base or frame. A large nut at the top of the center pin furnishes the means of holding the turntable down and in the proper position relative to the crawler. (See sketch A & B.)
- (2) In the other method, no center pin is used and the turntable is held in position by hook rollers which ride on the bottom of the roller path and act like clamps to hold the turntable down and anchored to the crawler. (See sketch C.)

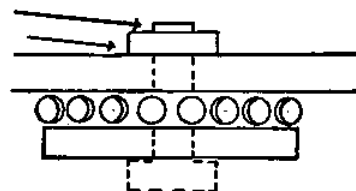


SKETCH A

One method is to employ 4 equally spaced rollers.

This is known as:

PLAIN ROLLERS

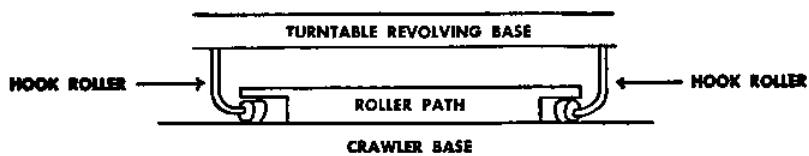


SKETCH B

This is similar to sketch A except that rollers are used completely around the circumference of the roller path.

This is known as:

CIRCLE OR NEST OF ROLLERS



SKETCH C

interviews, but we have included in these instructions some diagrammatic sketches which will help you understand the meaning of all terms and questions. Do not forget that many terms which may be Greek to you are part of the everyday language of all respondents. However, carry these instructions with you so you may refer to them constantly, and even show them to the respondent if need arises.

Since these diagrammatic sketches and the descriptions accompanying them may be helpful in making clear other industrial research problems, we are including them here as Figs. 1, 2, 3 and 4.

Questionnaire Sections. As mentioned above, our questionnaire was divided into eight sections or subjects:

- A—Identification (7 questions)
- B—Basic Factors and Uses (3 questions)
- C—Revolving Turntable (15 questions)
- D—Crawlers (12 questions)
- E—Booms (11 questions)
- F—Price (1 question)
- G—Accessories (1 question)
- H—Manufacturers (1 question)

Some typical questions. An indication of the kind of questions may be helpful on other industrial surveys. All questions were constructed so that their answers could be clearly and accurately reported and later tabulated by means of Hollerith punched cards. In studying these questions, remember Charles Kettering's statement: "A problem clearly understood is generally fairly simple."

Under Section E—Booms, our first four questions were as follows (note that all answers may be correctly recorded and reported by a simple "X" in the proper place, and that all answers are precoded for the card-punching operation):

E—BOOMS

1. What is the minimum acceptable length of crane for $\frac{3}{8}$ -yard and $\frac{1}{2}$ -yard machines:

$\frac{3}{8}$ -yard		$\frac{1}{2}$ -yard
20 feet		20 feet
25 feet		25 feet
28 feet		28 feet
30 feet		30 feet
47-1	-2	48-1
-3	-4	-2
-3	-4	-3
-4	-4	-4

54 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

2. What type of crane boom is preferred:

a. Channel type		49-1
Lattice type		-2
b. One piece type		-3
Two piece type		-4

3. When two-piece type of crane boom is used, what method of inserting center sections is preferred:

Bolted connecting plates		50-1
Pin connections		-2
Flange connections		-3

4. What are the *minimum* lifting capacities acceptable to you at 10' and 25' radius on $\frac{3}{8}$ -yard and $\frac{1}{2}$ -yard units:

At 10' Radius		$\frac{3}{8}$ -yard Unit	At 25' Radius	
8,000 lbs.		51-1	1,500 lbs.	52-1
9,000 lbs.		-2	2,000 lbs.	-2
10,000 lbs.		-3	2,500 lbs.	-3
11,000 lbs.		-4	3,000 lbs.	-4
12,000 lbs.		-5		
		$\frac{1}{2}$ -yard Unit		
10,000 lbs.		53-1	3,000 lbs.	54-1
11,000 lbs.		-2	3,500 lbs.	-2
12,000 lbs.		-3	4,000 lbs.	-3
13,000 lbs.		-4	4,500 lbs.	-4
14,000 lbs.		-5		

These four questions are typical of all those asked, and each was designed both to bring out the definite answers of all respondents and to give the client the range of requirements in his designs.

Making the question clearly understood. The client needed to know the specific range of the dimensions to use in his new models. Obviously (while it would be easy to ask innumerable questions to obtain these data), we faced the problems of recording correctly the answers given by respondents and the tabulation of such answers. For example, under Booms, we wanted to know the highest bank on which the respondent would use the $\frac{3}{8}$ -yard and $\frac{1}{2}$ -yard units to load trucks. A simple question to this effect was easy to phrase and answer. But it had to be followed by four dimensions and ranges for each size of unit which would be acceptable on the highest bank named by the respondent. This appeared to be not one question, but eight. But once clearly understood, it became quite simple. By use of diagrammatic sketches it was made entirely clear to the inspectors, who in turn were able to make it equally clear to the respondents by simply showing them the questionnaire.

Again, since this method may prove useful on other technical industrial surveys, Question #10 from Section E is reproduced in its entirety.

Favorable response to study. As the inspectors began to submit their reports from the field—the total was 425—it could be seen that they reflected keen interest by practically all respondents. From Birmingham the inspector wrote: "Most of the questions asked in the survey were not familiar to engineers and contractors. All of them pointed out that these could be answered more authoritatively by the operator of these shovels. In fact, all of them wanted to have their operators interviewed and then endorse the result. On the other hand, the operator who was interviewed very readily answered each question and gave the impression of intimate knowledge and familiarity with each phase of the survey."

Another inspector from Phoenix added to one report: "Mr. _____ appeared to be highly honored and pleased that he should be selected as respondent in this survey." Rarely do market research men meet with this attitude in consumer research.

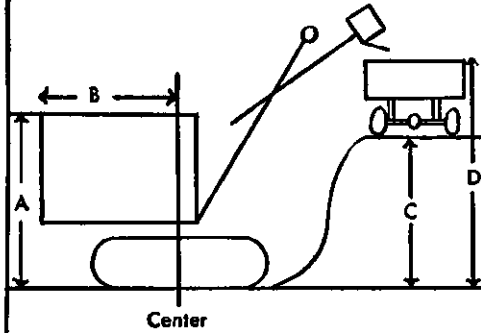
The most successful research surveys are those giving something to the respondent as well as taking something from him. This survey was a good example of mutual interest. This is well stated in the words of our inspector in Little Rock: "The men interviewed were very willing to cooperate in giving answers to these questions, and they all seemed to have been somewhat interested in the questionnaire as it did bring to light information on certain parts of the machine that they had not heretofore thought of."

For good and sufficient reasons the client did not permit his name to be divulged to any respondent and of course the field staff was kept in ignorance of it. This was regrettable, for most well-conducted market research surveys can be made makers of considerable goodwill for their sponsors. As an example, a Los Angeles inspector noted on one of his reports: "The respondent

10. Fill in dimensions and ranges in sketches below which would be acceptable to you for $\frac{3}{4}$ yd. and $\frac{1}{2}$ yd. machines:

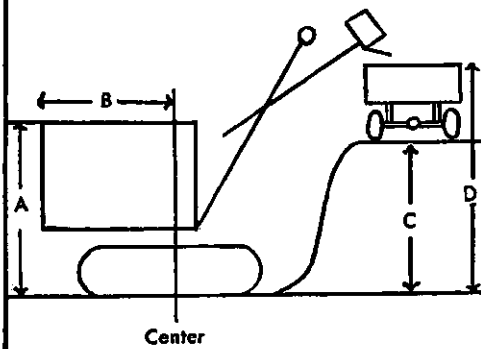
$\frac{3}{4}$ yard

- A. _____ 62-
- B. _____ 63-
- C. _____ 64-
- D. _____ 65-



$\frac{1}{2}$ yard

- A. _____ 66-
- B. _____ 67-
- C. _____ 68-
- D. _____ 69-



stated that in his opinion a survey of this type was something that had been lacking in manufacturers' production procedure for many years, not only in the manufacture of the subject equipment, but also of various other manufactured products that are in every-day use by business men, laborers, housewives, and the public in general."

Results for the sponsor. With our great conversion to war production, the client could not measure the results of this study as he would under more normal conditions. However, he did immediately express his pleasure with the results when they were submitted to him, and shortly thereafter reported that the survey results had practically ended all their technical discussions. All their engineers had to accept the evidence obtained from their market. He mentioned that our study had proved they had wide dimensional margins on many points on which they thought they had narrow limits. Other dimensions which had been presumed to have large ranges were found to be exceedingly restricted.

Industrial research in this country has not progressed as far or as rapidly as has other consumer research. This may be due to the complicated and technical problems involved. It is hoped that this discussion of one such study may open the way for many more uses of known market research techniques to the interesting industrial field.

Chapter V

THE MOVEMENT OF BRANDED GOODS TO THE CONSUMER

SAMUEL G. BARTON
Industrial Surveys Company

THE greatest progress toward the scientific approach to problems of efficient marketing of consumers' goods has probably occurred in the past ten years.

The goal sought is the ability to evaluate people accurately, individually, as consumers of any type of commodity and then to allocate the form and intensity of merchandising efforts to conform to the value and characteristics of the individual consumers as customers for the specific commodity.

In the simplest types of marketing this task is easily solved. The tailor who personally knows his customers selects his cloth with a particular customer in mind. In selling he varies his sales talk to suit the individual purchaser—emphasizes durability with one; style with another; and low price with a third. And finally, the suit is cut, made, and trimmed to the particular dimensions and whims of the individual customer. That individual company which applies the greatest skill in fitting its efforts to the market pattern is the one which will establish the greatest competitive gains and accordingly—the greatest real gains in profits.

The degree to which all manufacturers become more efficient in marketing directly affects both our rate of economic progress and the achievement of reducing the wide economic swings of depression and inflation.

In the field of consumers' goods distribution there are generally four levels at which progress or sales may be studied.

1. Manufacturer shipments to wholesalers.
2. Wholesaler receipts, inventories and shipments to retailers.
3. Retailer receipts, inventories, and sales to consumers.
4. Consumer purchases, inventories, and consumption.

Of the 10 million dollars spent each year for facts and figures, the great bulk (6 or 7 million) is spent in research at the consumer or final level.

Only at the consumer level can people be individually evaluated as customers, then divided into more or less homogeneous groups, which in turn can be identified and labeled in terms the manufacturer can use in allocating the form and intensity of his merchandising effort.

An individual manufacturer's sales figures will crudely estimate his consumer sales on a relatively long-term basis, but such figures, of course, show nothing about the movement of the total market including his competitors' sales.

Wholesale or retail sales can *theoretically* report *total market* movement. The total movement, however, can be broken down in terms of *consumers* on only two factors, *geographical location* and *type of outlet*, and for these two, only by inaccurate indirection.

In practical research application, wholesaler and retailer sales figures must be rejected for studying many of the major problems because of two considerations:

1. Such figures are *bulk* figures and do not permit direct identification of an individual sale to a specific kind of consumer. This will be discussed more fully.
2. In addition to this schematic limitation, such "bulk" figures are also inaccurate as such because of the growth of *multiple* channels of distribution including mail order and direct house to house or wagon sales, which forms of distribution are not susceptible to the research technique of estimating sales through the co-operation of a sample of retailers.

Aside from technical limitations, bulk figures obtained by checking sales of retail stores are made further inaccurate by the practical limitations imposed by major chains or multiple stores. In the United States these major chain stores do not permit commercial researchers to enter their stores for the purpose of estimating their sales. Such chains include the major syndicate, or "5- and 10-cent stores," and the Atlantic and Pacific Tea Company whose outlets blanket the most heavily populated half of the United States. Others with the same policy include the Kroger Stores, Safeway stores, and several of the important department store and drug chains.

Most of these major chain outlets have private brands sold only by their own stores and many of these brands actually rank first or are among the top four or five in sales in the total market.

At present, figures drawn from a sample of drugstores will represent as little as *one-half* of the total sales of many important "drugstore type" commodities. Likewise, with the impossibility of sampling the large food chains, any market estimate drawn solely from independent or small multiple food

outlets cannot accurately estimate the absolute total or even the relative competitive movement of food products. In the United States this is particularly true of coffee, tea, flour, soap, cereals, dentifrices, toothbrushes, vitamins, etc., but, to a greater or lesser extent, this limitation contributes to serious error in the measurement of all types of consumers' goods.

The errors caused by the above deficiencies are of the *same type* as the historical research error made by the *Literary Digest*. Unlike the poll, however, manufacturers have had no presidential election to make obvious to all the error of incomplete and unrepresentative sampling of the total market.

During the past six years Industrial Surveys Company has devoted its primary efforts toward measurements at the consumer level. One of the reasons for this can be quickly observed in the following chart showing the brand distribution of toothbrushes.

During 1943 it was the general impression of manufacturers, based on a sample of sales of drugstores, that Dr. West held approximately one-third of the toothbrush market, with Tek holding about one-fifth; Prophylactic, one-seventh; and all other brands, about one-third.

Panel results showed this estimate to be far from the true brand distribution. It is, however, a fairly accurate picture of toothbrush sales among drugstores. This is obvious substantiation of the statistical fact that nothing about the whole can be represented by sampling a disproportionate part.

In view of the above technical and practical limitations of applying bulk figures drawn from the manufacturer wholesale or retail levels to the problems of measuring the movement of branded goods to the consumer, industry has turned to various devices for studying the problem directly at the consumer level. Research efficiency requires that extensive corollary marketing data be obtained in the process of developing accurate estimates of over-all commodity and brand movements.

MEASURING THE MOVEMENT OF GOODS AT THE CONSUMER LEVEL

To be most accurate, any device for measuring the movement¹ of branded goods to the consumer must have several attributes:

It must provide a quantitative basis of measure; pounds, units, dollar volume, number of purchase transactions as well as number of families buying, number of people buying.

The approach of the Industrial Surveys Company to this problem of measurement of consumer movement is called "The National Consumer Panel."

When used by intelligent management, the panel becomes the most accurate

¹ A consumer inventory of consumption goods is not valid technique because no weighting is given to rate of consumption for each item.

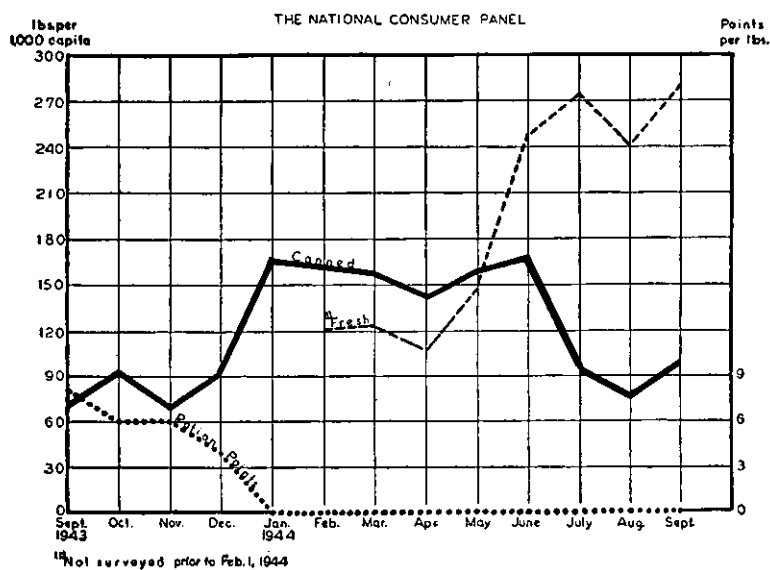
THE MOVEMENT OF BRANDED GOODS TO THE CONSUMER 61

and evidently the only device by which we may accurately study total retail commodity movement.

Aside from the many additional values which may be derived, we have reasoned that the panel technique, with an average 97 per cent accuracy in measuring the total market for consumers' goods, is better than techniques theoretically technically perfect which omit any estimate of movement of from 30 to 50 per cent of the total market for various commodities.

CHART I

UNITED STATES CIVILIAN PURCHASES OF
CANNED AND FRESH GREEN AND WAX BEANS



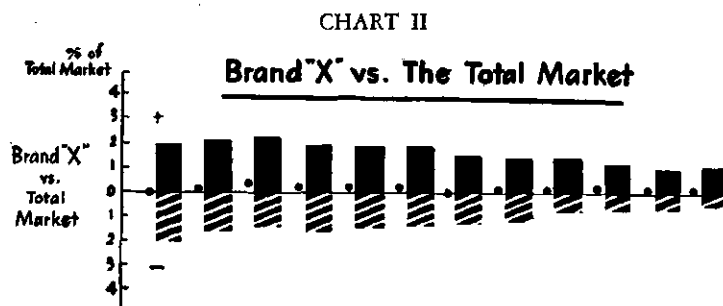
Although not yet well known in the recently developed field of marketing research, The National Consumer Panel was selected by the United States government to provide a quick and accurate index of the movement of rationed foods for wartime government control (see Chart I).

Other subscribers in the United States include the manufacturers and agencies for about 26 food and drug commodities. Additional companies are today revising their research, advertising and sales departments to make use of the National Consumer Panel.

In addition to measuring the over-all month-by-month or week-by-week brand sales through all types of outlets, the panel permits of accurate measure-

ment by observation of the primary causes for sales increases or decreases and study of the efficiency of merchandising and advertising methods.

Chart II (Brand Interchange) shows not only how much Brand X gained in sales, but the more causal picture of how much business was gained in gross and how much business was lost in gross between the two periods to produce the actual net change.



Illustrated in the above chart is the first stage of a continuing analysis of brand interchange. The solid black bar above the line shows the amount of business Brand X gained from all competition between each pair of periods. The striped portion of the bar below the zero line shows the amount of business Brand X lost to all competition for each pair of periods, with the black dots indicating the net change and its share of market between each pair of periods.

This, first of all, tells the manufacturer whether the customers of his brand are a large group of infrequent buyers, relatively disloyal customers, supported mainly by advertising or whether his market is tightly packed among a smaller number of highly loyal and steady users.

Such knowledge sets the basic merchandising and advertising policy of the manufacturer for the ensuing six months or year period.

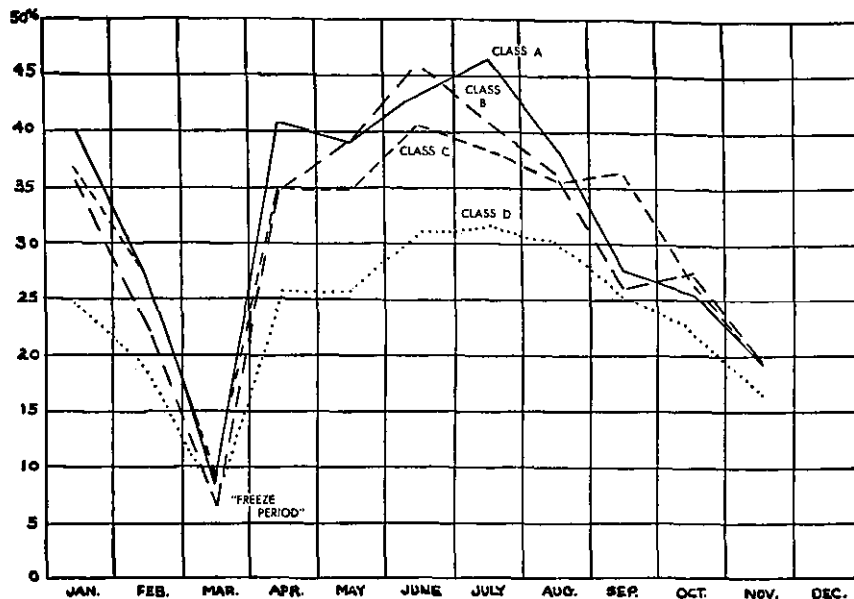
This knowledge of the market can be further refined. The National Consumer Panel can subdivide the gains by types of products and by individual brands. It can subdivide the losses by brands and accordingly can measure the net interchange of business between any one brand and any other (Brand Interchange Chart).

By this means the manufacturer can tell from last year to this how much business was gained or lost to competing brands B, C, D, etc. He may also measure the amount of business gained or lost due to improvement or deterioration of consumers as customers for the entire commodity: changes due to causes other than competitive shifts.

It will be obvious to the sales manager that, with the facility for focusing gains and losses on individual brands, it is possible to analyze the individual competitor's product, packaging, price, methods of distribution and advertising, and to take corrective measures.

With equal facility we can isolate those individual customers who reduced their buying of a given brand and study their family or personal characteristics. We can examine this "lost group" of customers from the standpoint of location, age, income, family size, and other attributes. We can also study their habits, such as type of product usually bought, price paid, size of package

Proportion of Families Buying Canned Meat
by Economic Class
Source: THE MAGAZINE PANEL-1943



bought, place of purchase, etc. In addition, we may study these lost customers from the standpoint of the nature and amount of our advertising which reached them and whether or not they were adequately reached by the manufacturer's special deals, introductory offers, or other forms of special merchandising effort.

Before examining the other two broad applications of the National Consumer Panel we should obtain a more complete picture of the technique itself. This paper, however, will not attempt to report all the experiments, tests of accuracy, and technical procedures which have built the method to its present high level of technical accuracy.

Two panels are operated by Industrial Surveys Company on a national basis—one continuously for food, drug, and media measurements; the other repet-

itive but noncontinuous for measurement of opinion—primarily opinion on consumers' goods, conditions of usage, methods of use, and opinion of new products. In addition, the company has panels that are operated in twelve local market areas. One operates continuously, the others are periodically continuous and are used for local media analysis and test market purposes.²

Panel technique means that the facts about attributes, purchases, and media exposure are collected repeatedly from the same body of families or people. This technique makes it possible to study the purchases of a family made *during one period* of time in relationship to the purchases they made during a previous period or to study them in the light of the varying volumes of advertising to which they were exposed in the past.

If all persons had unlimited memories, the panel technique would be unnecessary. It would be possible to ask a family, "What proportion (to just one decimal) of your soap did you buy from Ivory during each month of 1943? How many issues of magazines carrying Ivory advertisements came into your home during 1942? 1943?"

While everyone will agree that no consumer could answer such questions, the answers are requisite if we are accurately and thoroughly to study consumers' goods and the efficiency of merchandising efforts.

The operation of the various panels is fundamentally the same. The facts as to purchases of food, drugs, receipts of magazines and newspapers, and listening to radio programs must be kept daily and continuously in diary form in the home.

This work requires, on the average, three minutes per day for the participating family. It is an essential factor and a critical requirement that the family diaries be collected each week. After preliminary instruction, making the entries properly is far simpler than filling out a Sears, Roebuck mail-order blank or a pre-New Deal tax return. Because the time required averages three minutes per day and because the work itself is interesting, some families, we have found, would be willing to participate for a period of time without compensation.

To obtain a cross section and to *maintain continuity* of reporting among *all types* of families, however, it is necessary that each family be compensated for their work. This is only fair, since the accuracy of the reports is dependent upon the co-operation and work of the family in the privacy of their own home.

Compensation is paid the family in the form of points, which are credited each week. The family may then select from a premium catalogue an article of merchandise each month, or, if they prefer, points may be accumulated to

² Apart from the National Consumer Panel, as of April 1946, Test Market Panels are operated continuously in nine cities.

purchase large items such as furniture, watches, luggage, etc., at Christmas-time or at a birth or wedding anniversary.

The degree to which our efforts have been successful in developing a high technical accuracy has been established not only in many field tests, but has been proved in our estimates of an extensive list of individual manufacturer brands' sales where extensions of the panel reports have regularly agreed within a few pounds per capita with known industry and brand sales for one and two year periods (see Chart III).

CHART III

Panel Reports

Extensions to U.S. Totals	Variance from Known Mfg. Sales Basis - One Year
Hi-frequency Product Classes	± 2%
Individual Brands	± 6%
Low frequency Product Classes	± 3%
Individual Brands	± 10%

**Proven on. Tooth Brushes, Dentifrices, Cold Cereals,
All Purpose Flour, Coffee**

The above table represents a summary of valid comparisons available between manufacturers' estimates of their yearly sales to consumers and projections of the panel. An opportunity for an outstanding comparison between panel and actual sale was derived from one regional manufacturer of premium flour. Since sales of this premium product to industry and institutions are practically nil, a valid comparison over an extended time period was available. A total of two years' manufacturer's shipments to retailers came within three-tenths of 1 per cent with the panel. Similar high degrees of agreement have been established in other instances of valid comparison for both regional and national brands.

Our procedures in developing the technique to its present accuracy level were reported before the American Marketing Association meeting devoted to panel research on June 9, 1944, at Philadelphia.

Leaving technical matters, we shall examine the two additional broad applications which were earlier mentioned.

The National Consumer Panel obtains continuous records of radio program listening, magazine and newspaper receipts. The listening to specific programs and receipt of specific issues is recorded. As will be recalled, at the close of any period we have a complete record of the purchases of any given commodity. The total purchases of an individual family for an entire year may

then be tabulated. We may then divide all families into subgroups according to the *amount of the commodity they purchased during the year*. We may then determine what proportion of each of these heavy, medium, light users, or non-users were reached by the advertising of a given brand during the year. We may examine the result to see if the broadest coverage is among the highest buying rate families with coverage diminished in proportion to lower buying rate. Further study is then made as to the distribution of the advertising schedule to learn whether the greatest contact frequency is also among the highest buying rate group and lowest frequency among the poorer market groups.

Should we find (as we frequently do) that the greatest market coverage and highest contact frequency are not among the highest buying rate families, the structure of media schedule may be rebuilt element by element to produce the desired degree of fitness to the market in accordance with the advertiser's policy. (In certain stages of the development of a new commodity such as vitamins, frozen foods, dehydrated soups, etc., it may be a valid part of the advertiser's plan to allocate a portion of the schedule toward the development of new markets.)

It is apparent that the panel is useful, not only in the development of broad media and advertising policy as to division between radio, printed space, etc., but in increasing the ability of the time and space buyer in the selection of the individual components of the media schedule.

In actual operation we determine for any existing (or proposed) schedule, the number of contacts, based on the previous year, which the schedule actually made with each dollar's worth of commodity purchases. The assumption is that the characteristics of circulation, coverage, and penetration of the media will remain basically the same during the ensuing 6- or 12-month period.

Individual media are evaluated in the same fashion. A short cut in the preliminary selection of media has been developed which permits of selection on the basis of relative efficiency of any individual media to reach various subdivisions of the market for an individual commodity.

The second broad application, derived solely from the panel, is the ability to study the continuing performance of consumers subdivided according to their number of exposures to advertising or the exposure to special merchandising efforts such as deals, special offers, and the like.

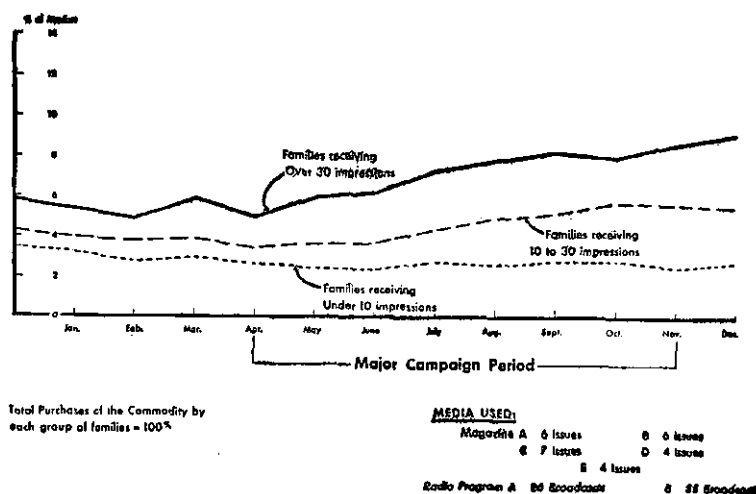
This means evaluating the gross sales effectiveness of advertising and effectiveness of special merchandising.

To demonstrate this application we shall examine one illustration of the advertising effort of a soap company (Advertising Sales Effectiveness Chart—Chart IV).

To construct this analysis, all families were divided according to the number of contacts of the brand's advertising which they received during the campaign period.

CHART IV SALES EFFECTIVENESS OF ADVERTISING

Brand 'X' Share of The Market
Trend by Number of Magazine & Radio Impressions



This chart illustrates the most feasible technique yet developed for estimating the volume of business produced from advertising. The actual evaluation is made on a tabular basis by comparing the gains per thousand families between the exposed groups and the unexposed groups.

On this chart, the proportion of their total soap purchases made from the advertised brand are plotted for three months preceding the campaign, for the six months of the campaign, and for three months following.

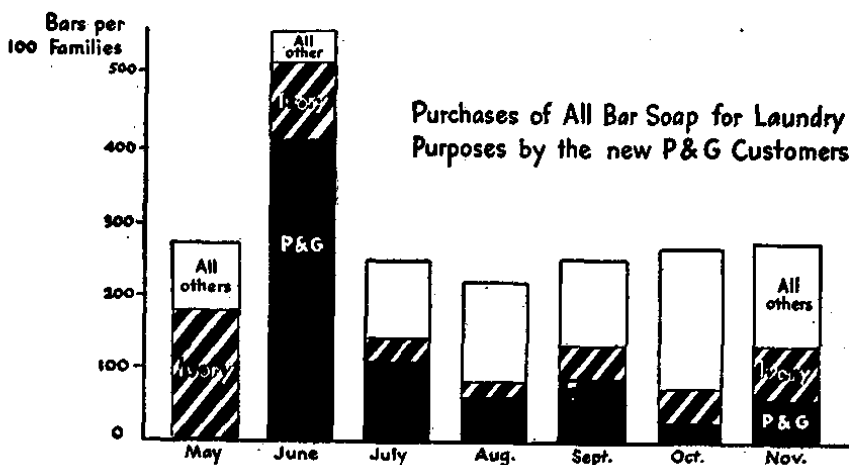
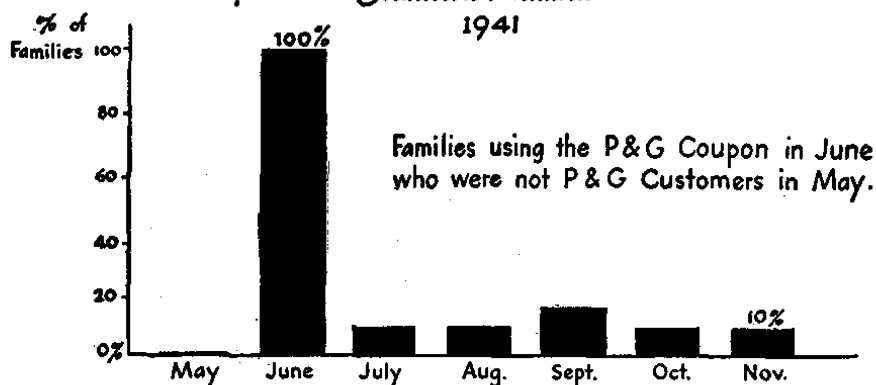
No significance is attached to the differences between the advertised and unadvertised groups at any period of time. The sales effectiveness is measured by the amount of improvement within the advertised group as customers of the brand compared to the amount of improvement found within the unadvertised group.

The panel technique can be readily applied to the study of the relative change found among consumers exposed to a deal or introductory offer as compared to the trend of purchases found among the group not so exposed (See Chart V).

CHART V

The Consumer Panel Evaluates A Deal

P&G Coupon offer on 6 Bars
Evansville Indiana
1941



This chart illustrates one technique of evaluating the sales produced by special merchandising effort. While the deal was successful in introducing new customers to P. & G. Naptha Soap, gains made were primarily at the expense of other brands distributed by the same company. The above data must be construed solely as a demonstration of technique rather than any representation of performance of the total market because of inadequate sample and inadequate time series preceding the coupon deal.

While the validity of this technique is now well established in scientific experiments, its application is limited by the complexity of the problems and the size of the panel sample. The degree of refinement of results is also controlled by the frequency with which the commodity and the brand are purchased.

Thus a highly refined evaluation may be made of a major brand of coffee because of the low relative variance. The analysis may frequently be carried to the point of separately evaluating individual components of the advertising structure. At the other extreme would be the problem of evaluating the sales resultant from a small campaign split among many media forms by a minor brand of a low consumption rate commodity such as Worcestershire Sauce.

The National Consumer Panel is based upon the participation of 3,000 families³ representing slightly more than 9,000 persons. While it is being regularly increased in size, it would need to grow manyfold before it can be used to evaluate minor campaigns. In a relatively "ideal" case, we are able to say media A produced \$1,000,000 of additional business plus or minus \$100,000 while media B produced \$1,500,000 business plus or minus \$50,000 establishing a greater selling power for media A (cost being equal).

In cases of complex and small campaigns in low-frequency commodities, the tabulations are more likely to show that all media combined produced, for example, \$100,000 of additional business plus or minus \$200,000. The high apparent variance evidently is due to the interaction of unmeasured secondary influences, and to the interaction of the advertising and economic level, educational level, and other attributes.

While complex analyses of variance can probably refine the evaluations, it has been found that the major requirement is an extended time series of purchasing and advertising.

For the purpose of estimating sales effectiveness in consumption goods, inventories of the amount of a given brand in homes at any given time or even on repetitive measurements is useless because of the lack of correlation between inventory and actual rate of consumption. Effectiveness must be based on measurement of gains in quantitative rate of consumption. The National Consumer Panel, in recognition of this requirement, is designed to obtain records of consumption and exposure from families on a consecutive four-year basis.

The entire panel accordingly will be rotated every five years. This means that 20 per cent of the *average* sample size will provide reports for the full four-year period. At the end of three years (Sept. 30, 1944) it is apparent that this objective has been realized. The present *experimental panel* began with approximately 2,000 families. This means that after sample adjustment we

³ Four thousand families as of April 1946.

shall probably have 500 families on which to base the four-year analysis. While perhaps adequate for much of our *experimental* work, actual measurements of sales effectiveness for major commodities and advertising campaigns require a *static sample* of several thousand families. This in turn, of course, means an average sample approximately four times the size of the four-year static sample.

Until this larger commercial sample is realized, the greatest gains in increase of advertising efficiency are being derived from achieving a better fitness of the advertising to the market as earlier described in this chapter. This procedure can be carried out with much shorter term static samples and therefore with smaller total samples involving 2,000 to 5,000 families.

In review, through the panel technique we know the following things about each purchase transaction and each preceding and succeeding purchase.

Day of Purchase	Type, color, flavor, etc.	Place of Purchase
Date of Purchase	Size of Package	Person Buying
Brand Purchased	Number of Units	Age, sex, occupation, etc.
	Price	
	Deal or Special Offer	

In addition, we know the following things about each consumer family (obviously only a partial list):

Location	Size	Number of Advertising
Economic Class	Age and Sex	Exposures to Radio,
Education of Adults	Occupation and Income	Magazines and Newspapers
	of Members	

Each of the above facts, as they are collected each week, is punched into Hollerith cards. With modern electrical equipment, including alphabetical tabulators, collators, etc., we have developed the methods by which each of the above facts about one purchase can be related to each other purchase and in turn to each continuing fact that we learn about the family—its attributes and its exposure to advertising and sales efforts.

The ways in which these facts can be combined are astronomical in proportion—almost equal to the varied and complex problems which marketing, sales, and advertising men are called upon to solve.

With the help that we have received in our research efforts from the United States and British governments and the research and statistics staffs of the manufacturers of foods and drugs, we are hopeful and confident that the National Consumer Panel will continue to contribute to progress toward more understanding of the consumer market.

Chapter VI

PSYCHOLOGICAL BRAND BAROMETER

ALBERT D. FREIBERG
The Psychological Corporation

INTRODUCTION

THE nature of the Psychological Brand Barometer can best be illustrated by some actual results. The table on the next page shows results of the February (one of four each year) Coffee Barometer in four consecutive years. In each study the question "What brand of coffee did you buy last?" was asked of housewives in 55 cities and towns of the United States.

The Psychological Brand Barometer illustrated by this table of four studies is a measure of sales not in the economic, but in the psychological sense. It measures not in terms of pounds of coffee, or of dollar volume of coffee bought, but in terms of people's behavior. It is a measure of customers, a report of what brands families use. Its purpose is to reveal whether, from study to study, more or less families are buying a specific brand and to show such trends in relation to competing brands.

ECONOMIC VS. PSYCHOLOGICAL DATA

Increases or decreases in the number of families or individuals using a particular product *may* or *may not* parallel fairly closely increases or losses in volume, in terms of either units or dollars. The sales of a certain brand may be falling off even while the proportion of customers for that brand is increasing. The reverse may also be true. These barometers were designed to measure changes or trends in customers; if at the same time they give some indication of volume, then that is an additional factor in that particular case. A study of differences between customer trends as measured by these barometers and the usual sales report trends indicates that each measures a relatively indepen-

Brands	COFFEES			
	STUDIES MADE FIRST WEEK IN FEBRUARY			
	1941 %	1942 %	1943 %	1944 %
Maxwell House	16.7	15.4	18.2	18.2
Chase and Sanborn	8.9	8.0	9.5	11.0
Bokar (A. & P.)	3.5	3.2	5.7	4.0
Eight O'Clock (A. & P.)	10.7	11.7	6.1	5.6
Red Circle (A. & P.)	4.0	5.0	3.5	3.3
Hills Brothers	6.5	7.5	8.0	8.3
Beech-Nut	4.0	3.0	3.2	4.2
Jewell Tea Company	1.2	.9	.9	.7
Folgar	5.5	5.7	8.0	7.3
Sanka	3.4	3.1	2.6	2.8
M. J. B.	.6	.6	.7	.5
Del Monte	.7	1.1	.8	1.1
Boscul	1.3	1.0	1.7	2.4
White House	.3	.2	.5	.2
American Stores (Asco)	1.1	1.6	1.3	1.4
Mannings	.4	.3	.2	.2
La Tourraine	.6	.9	.5	.4
Kybo	.8	.4	1.0	.8
Yuban	.1	*	*	*
Miscellaneous	29.7	30.4	27.6	27.6
Total Per Cent	100.0	100.0	100.0	100.0
Total Users	9389	9383	4667	9607
Don't use	4.9	4.9	3.3	3.1
Don't know	1.2	1.3	2.8	.8
Total Interviews	10,000	10,000	5,000	10,000

*Less than .05 per cent.

dent set of facts. Usually fluctuations in the number of customers are more gradual than changes in sales volume.

Lack of comparability between trends in dollar or unit volume and numbers of customers may occur for a variety of reasons such as:

1. Special sales at reduced prices so that the same number or even fewer people buy as much or more of the product as was purchased in a comparable preceding period.
2. The anticipation by customers, wholesalers and dealers of shortages; for example, sugar shortages and coffee shortages.
3. Fear of rising prices which may cause dealers, families, and also wholesalers to buy large quantities ahead.
4. A general increase or decrease in buying power so that the same number of families may buy more or less of a particular product than they did before. During the past few years, for example, the so-called ill-fed, ill-clad, ill-housed one-third has acquired tremendously increased buying power and has consumed quantities of food, spent millions of dollars on clothing, etc., which three or four years ago would have been undreamed of.

PSYCHOLOGICAL BRAND BAROMETER

COFFEES†

SERIES 41

BASED ON 10,000 INTERVIEWS MADE FEBRUARY 1-6 IN 55 LOCALITIES.
REPORT COMPLETED MARCH 7, 1941.

	SOCIO-ECONOMIC GROUPS				Total %
	A %	B %	C %	D %	
Maxwell House	22.2	19.0	16.0	11.2	16.7
Chase & Sanborn	11.3	9.9	9.0	6.1	8.9
Bokar (A. & P.)	4.0	4.0	3.5	2.3	3.5
Eight O'Clock (A. & P.)	5.7	8.0	10.7	17.7	10.7
Red Circle (A. & P.)	3.2	3.5	4.4	4.3	4.0
Hills Brothers	6.4	7.7	6.0	5.8	6.5
Beech-Nut	6.3	5.7	3.6	1.3	4.0
Jewell Tea Company	.3	.8	1.3	1.7	1.2
Folger	5.8	6.0	5.6	4.6	5.5
Sanka	4.3	3.7	3.1	2.6	3.4
M. J. B.	.8	.7	.6	.5	.6
Del Monte	.5	.8	.9	.5	.7
Boscul	1.1	1.1	1.2	1.6	1.3
White House	.2	.3	.3	.3	.3
American Stores (Asco)	.4	.9	.9	2.2	1.1
Mannings	.8	.3	.5	.3	.4
La Touraine	.5	.4	.6	.9	.6
Kybo	.3	.7	1.1	.6	.8
Yuban	.1	.1	*	*	.1
Miscellaneous	25.8	26.4	30.7	35.5	29.7
Total %	100.0	100.0	100.0	100.0	100.0
Total Users	967	2880	3763	1779	9389
Don't Use Coffee	2.0	3.2	4.8	9.0	4.9
Don't Know	1.3	.7	1.1	2.0	1.2
Total Interviews	1000	3000	4000	2000	10,000

Question: What brand of coffee did you buy last?

†This table copyrighted by the Psychological Corporation.

*Less than .05%.

5. Seasonal changes in volume may not at all correspond to trends in number of customers.

6. The rationing rules and regulations have brought about many radical changes in the relationships between customers and volume. They have also produced changes in customers' buying and consuming habits, not only in regard to quantities but also in regard to brands and types of products so that some of the usual significance of sales figures has to a large extent disappeared.

These and similar factors affecting figures on sales volume tend to obscure the critical element in many sales and advertising programs, namely, whether or not new customers are being obtained. Perhaps of equal importance is that these factors obscure what is happening to competitors' brands. Obviously, it

74 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

is important for a company to know not only how well it is doing, but also how well a competing company is doing. The Brand Barometers give such a comparative picture in terms of the number of customers. Results of customer changes are shown not only for a particular brand, but also in relation to all other competing brands. And customers are, basically, what companies try to get.

RESULTS BY SOCIO-ECONOMIC GROUPS

Results are always shown by four socio-economic groups: A, B, C, and D; A representing the highest and D the lowest. The following copy of the February, 1941, Coffee Barometer illustrates some significant differences in brand dominance by socio-economic groups.

COFFEES†					
SERIES 53					
BASED ON 10,000 INTERVIEWS MADE FEBRUARY 1-14 IN 55 LOCALITIES.					
REPORT COMPLETED MARCH 13, 1944.					
SOCIO-ECONOMIC GROUPS					
	A	B	C	D	Total
	%	%	%	%	%
Maxwell House	18.8	19.0	17.8	17.3	18.2
Chase & Sanborn	9.7	11.0	11.0	11.6	11.0
Bokar (A. & P.)	3.8	4.5	3.8	3.9	4.0
Eight O'Clock (A. & P.)	2.7	3.9	5.9	9.4	5.6
Red Circle (A. & P.)	2.5	3.6	3.3	3.2	3.3
Hills Brothers	9.3	8.7	8.3	6.9	8.3
Beech-Nut	7.9	5.4	3.9	1.2	4.2
Jewell Tea Company	.4	.4	1.0	1.0	.7
Folger	7.7	8.7	6.5	6.7	7.3
Sanka	3.6	2.6	2.9	2.3	2.8
M. J. B.	.5	.6	.4	.6	.5
Del Monte	1.3	1.5	.9	.6	1.1
Boscul	1.4	2.2	2.9	2.4	2.4
White House	.1	.1	.1	.3	.2
American Stores (Asco)	.6	.9	1.5	2.5	1.4
Mannings	.5	.2	.3	.1	.2
La Touraine	.1	.3	.3	1.0	.4
Kybo	.1	1.0	.8	.5	.8
Yuban	.1	*	.1	*	*
Miscellaneous	28.8	25.3	28.2	28.4	27.5
Total %	100.0	100.0	100.0	100.0	100.0
Total Users	972	2915	3861	1859	9607
Don't Use Coffee	2.0	2.2	2.7	6.0	3.1
Don't Know	.8	.6	.8	1.0	.8
Total Interviews	1000	3000	4000	2000	10,000

Question: What brand of coffee did you buy last?

†This table copyrighted by the Psychological Corporation.

*Less than .05%.

It will be observed immediately that the high- and medium-priced brands such as Maxwell House, Chase & Sanborn, and Beech-Nut show definite trends by socio-economic groups, being more popular in the higher groups. The less expensive brands (*e.g.*, Eight O'Clock and Miscellaneous) show greater acceptance in the lower groups.

Three years later many of these differences had disappeared under the influence of rationing, shortages, increased buying power (especially in the C and D groups) and other factors.

Maxwell House now shows only slight differences by socioeconomic groups. Chase & Sanborn has reversed its trend, now being slightly more popular in the C and D groups. Miscellaneous brands, many of which have limited distribution and are generally cheaper, now show no trend. Eight O'Clock shows differences by socioeconomic groups of only about half the magnitude of three years earlier. Such facts are useful to companies in the determination of their merchandising and advertising policies.

SEPARATE REPORTS FOR MEN AND WOMEN

Some products, such as shaving creams, are reported only for the man in the family. Other products, such as toilet soap, are reported separately for men and women. Still other products, such as coffee, are reported in terms of family use. The following abbreviated table from the November, 1941, Barometer shows, for example, a comparison of usage of principal brands of toilet soap by men and women:

TOILET SOAPS USED BY MEN AND WOMEN*

<i>Brands</i>	<i>Women</i> %	<i>Men</i> %
Lux	20.3	15.5
Palmolive	16.0	15.1
Ivory	15.8	18.8
Camay	10.5	7.9
Lifebuoy	6.0	15.4

*Based on 10,000 interviews made November 1-6 in 55 localities. Report completed December 3, 1941.

Question: What toilet soap did you buy last for yourself?

Differences of this kind are of interest not only to the sales and merchandising departments of the advertiser but also to the advertising department and to the advertising agency, especially the copy writer, who writes some copy specifically for men and other copy specifically for women.

EXPERIMENTAL ORIGIN AND HISTORICAL DEVELOPMENT

The Psychological Brand Barometer was started and developed by Henry C. Link of the Psychological Corporation. [1]¹ After preliminary and local experimental work, the first large study of 1,578 housewives was made in March, 1932. Experimental studies were continued during the remainder of 1932 and during 1933. In September, 1934, the first commercially sponsored full-sized Brand Barometer study of 4,000 interviews was made in 47 localities scattered throughout the United States. In this study all interviews were made in upper-middle income group homes, commonly called the B socioeconomic group. The B group was chosen because the experimental studies had shown it would indicate trends earlier than the population at large. Therefore, changes of buying habits by this group would serve as the first indicators of more general changes or trends.

In November, 1934, and every two months following this another 4,000 homes were interviewed. This schedule was maintained until May, 1936, when the first 10,000-interview study was made. The same cities and towns were used but all socio-economic groups were included. In July and September 4,000 interviews with the B group were made and in November 10,000 interviews in all groups were again made. This schedule of two studies of 10,000 interviews each, one in May and one in November, was followed until January, 1940. In 1940 all studies were based on 10,000 interviews and four studies were made each year, the first week in February, May, August, and November. By June, 1945, 68 Brand Barometer studies and 442,000 interviews had been made. This series of studies is the oldest continuous and current poll of buying behavior in the United States.

THE SAMPLE

Interviews are made in 10,000 homes in each study. In all 10,000 homes the housewife or an adult female is asked questions on the first part of the questionnaire which covers products used by the entire family or products used by the women in the family. In approximately half the homes the second part of the questionnaire is answered by the husband or by an adult male in the family. This means that there are interviews with about 15,000 people in 10,000 homes. By asking questions of men on the second half of the questionnaire it is possible to get a sex comparison on buying behavior, attitudes, etc. This gives additional flexibility and range to these studies.

Each study is now made in 110 cities and towns. The change in sample

¹ Numbers in brackets refer to notes at end of chapter.

from 55 to 110 cities was made in November, 1945. The table below shows the number of cities and the number of interviews made in each geographical area and compares the number of interviews with the urban population in these areas as given in the 1940 census.

Geographic Area	Cities #	Interviews		1940 Census Urban Population*	
		#	%	%	%
East	32	3700	37.0		37.1
South	29	2050	20.5		20.5
Midwest	32	3150	31.5		31.5
Far West	17	1100	11.0		11.0
Total	110	10,000	100		100

*Cities and towns 2500 and over.

The sample is divided into four socio-economic groups, as follows: an upper group, designated as A, consists of 10 per cent of the families; an upper-middle group, designated as B, contains 30 per cent; a lower-middle group, designated as C, contains 40 per cent; the lowest group, designated as D, contains 20 per cent of the families. The problem of stratification and classification of families by socio-economic status is, of course, worthy of a book by itself. This is hardly the place to go into it in detail, but some comment seems indicated.

In each town or city the local psychologist in charge of the field work classifies the families in his particular locality or localities on a comparative rather than an absolute basis. The highest 10 per cent in terms of type of home, occupation, educational level, etc., are listed as A. In one community this might represent an annual cash income of \$5,000 or more, whereas in another community it might represent an annual cash income of only \$3,000. Regardless of the absolute level, the upper 10 per cent is classified as A. The national sample of A homes is obtained by adding together the 10 per cents of each individual city. In like manner the D, or lowest 20 per cent, is determined and designated. The same applies for the B, or upper-middle 30 per cent, and the C, or lower-middle 40 per cent.

In each city or locality the psychologist in charge makes a map designating the city areas in terms of these four socioeconomic groups. Assignments of interviews are made with the aid of this map. It has been the custom of the Psychological Corporation ever since the inception of these studies to specify that assignments in each city be widely scattered. An assignment of 50, 100, or 200 interviews in a city would be made in from four to fifteen different areas of the city. This scatter is made with the help of the socioeconomic map and has achieved from the start a rough but effective combination of area and quota sampling. As an added safeguard in the measurement of trends, accurate records are kept of streets and blocks covered in each study, and in

successive studies adjacent streets or blocks or odd- and even-numbered homes are used.

FIELD FORCE AND FIELD SUPERVISION

In each city one or more local psychologists act in the capacity of research associate of the Psychological Corporation and supervise the field work. Each psychologist is responsible for obtaining and training interviewers and for the handling of all details in regard to them, including payment. Training materials, instructions, and other helps are of course provided by the home office of the Corporation.

In each study of 10,000 interviews, more than 400 interviewers do the field work, making an average of 25 interviews or less per interviewer. This spreading of the risk of interviewer bias, carelessness, tendency or risk of faking interviews, etc., has been an established policy of the Psychological Corporation from the beginning of these studies.

The work of each interviewer is checked by means of follow-up interviews. In each city the psychologist or an interviewer unknown to the original interviewers makes calls on two or more of the homes originally called on by each interviewer in order to establish the fact that an interview was properly made at that home with the right person, etc.

THE QUESTIONNAIRE

The questionnaire is mimeographed and limited in length to one side of an "8½ by 11" sheet of paper. This allows for the asking of approximately 20 (or at most 25) questions. The interview takes a maximum of 20 minutes, and usually only about 15 minutes to complete.

Special instructions to interviewers cover the specific questions of each questionnaire. These are written at the main office of the Psychological Corporation and are studied by each interviewer before making his interviews. In addition, each interviewer fills out all the questions on the interview itself before making any interviews in order to become familiar with the recording technique.

Products Covered. Some of the products reported for family use are:

Mayonnaise
Margarine
Coffee
Tea
Beer for use at home
Brand
Type of container

Beer consumed away from home
 Brand
 Type of container

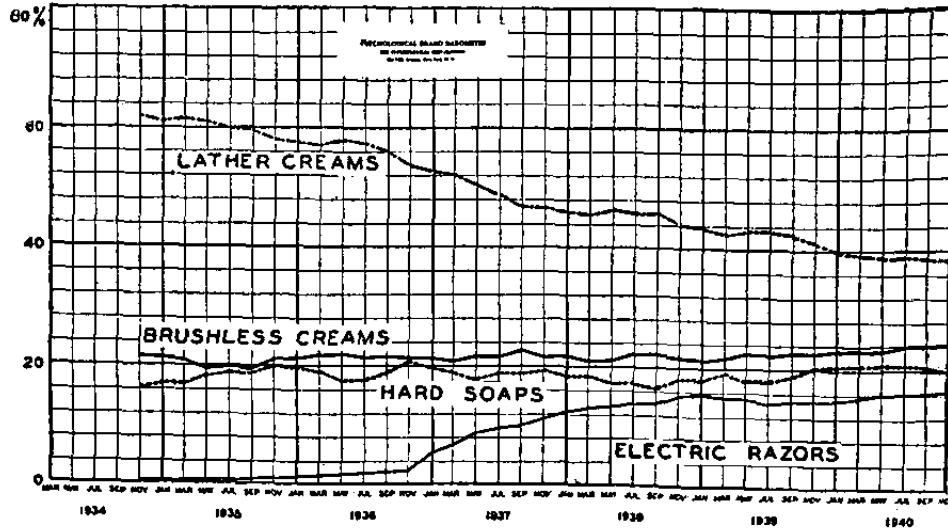
Some of the products reported separately for men and women are:

- Shaving cream for men
- Cigarettes for men and women
- Hand lotions for men and women
- Toilet soaps for men and women
- Shampoos for women
- Face creams for women
- Face powder for women

LONG-TERM TRENDS

The above and other products have been included fairly regularly in the Brand Barometer series since September, 1934. The following chart illustrates, for example, one of many long-time trends measured by these studies:

TRENDS IN SHAVING ACCESSORIES



Six-year Trends in Men's Shaving Habits. Above comparisons of men's use of shaving accessories based on 34 Brand Barometers of 128,000 interviews in the B income group. Special promotion has not helped to raise curves in this field.

This is certainly a dramatic picture of trends in men's shaving habits during a period of about seven years. The trend lines are trends of use, which is of particular significance in regard to electric razors, since it is well known

that there are more owners than users. For approximately one year the Brand Barometer indicated that less than 1 per cent of men in the upper-middle, or B, socio-economic group were using an electric razor. Then the percentage began to climb sharply until about the first part of 1939 when it leveled off and showed little increase. This does not mean, obviously, that no more electric razors were sold. In the first place, as some men discarded or gave away their electric razors, others bought. In addition, old models were traded in on new ones. There was, however, a static situation in regard to the net users. And this existed in the face of continued improvement in the quality of electric razors, a steady reduction in price, and some extremely high-pressure advertising campaigns and special promotions.

There was, however, another type of market expansion for electric razors. This occurred in an increase of users in the C and D economic groups as can be seen from the following table:

Socio-economic Status	PERCENTAGE OF MEN USING AN ELECTRIC RAZOR					
	1936		1937		1938	
	Nov. %	May %	Nov. %	May %	Nov. %	May %
C—Lower-Middle Group	.9	4.1	5.0	7.9	7.2	9.0
D—Lowest Group	.3	1.6	1.6	2.5	3.3	3.2

TIME SCHEDULE OF FIELD WORK AND REPORTING

The field work is completed during the first ten days of the months in which the studies are scheduled, that is, in February, May, August, and November. Much of the interviewing is done in the evening. Two weekends are allowed in order to ensure the best possible cross section of homes and the interviewing of available men in the family.

Completed interviews are sent to the New York office by first-class mail or by air mail. Upon arrival questionnaires are edited and inspected for stereotyped answers, for completeness, for accuracy with which the quotas have been completed, etc. Questionnaires are then coded and punched on Hollerith cards. The cards are then run and sorted to obtain the desired information, tables prepared, and reports with interpretations written. The reports are submitted to the clients in either typewritten or mimeographed form during the first week of the following month. This makes a total time period of approximately five weeks between the beginning of the field work and the presentation of the report to the client. This rapidity of reporting ensures that trends or findings of timely interest are up to date and current rather than historical in character.

OTHER USES FOR THESE PERIODIC STUDIES

In addition to the measurement of brand changes, these studies can be used for measuring the effect of advertising more directly. For example, the Triple Associates technique, described in Chapter IX in this book, can be used with a maximum of 5,000 interviews with men and 5,000 interviews with women, or with a maximum of 10,000 interviews with women. The reader will note that this technique consists in the asking of a question such as "What company advertises 'More Things for More People at Less Cost'?" Another advertising test which can be used is the Aided Recall Test. This type of test is illustrated by the questions: "What brand of hand cream or lotion have you seen or heard advertised lately?" and "What did the advertising say?"

Information about ownership of capital goods, such as automobiles; homes—kind and type, owned or rented; kind of heating or furnace, etc., can easily be obtained.

Attitudes on current topics, such as rationing, labor and industry, and similar types of opinions can be measured.

In short, this representative sample of 10,000 urban homes can be used for any of the purposes to which a repetitive series of studies is helpful or worth while.

EMPIRICAL DEMONSTRATION OF LAWS OF CHANCE

In April, 1934, the Psychological Corporation made a special Brand Barometer study of 5,165 interviews [2]. This study was used to make an empirical test of whether market research data of this kind conformed to the familiar Bernoullian type of distribution. Answers to questions were coded and punched on Hollerith cards. The cards were then sorted in random order into 51 sets of 100 cards each, each set of 100 containing the same proportion of interviews from each of the four geographical areas as were in the entire sample of 5,165. These 51 sets of 100 were then run on a variety of Brand Barometer questions.

One of the questions was on brands of tooth paste used. The following table from the original article shows the scatter of responses for two brands of tooth paste—brand A and brand E.

Brand A has an average of 16.22 per cent. The standard deviation of 16.22 per cent, as computed by the regular formula $(S.D. = \sqrt{\frac{pq}{n}})$ is 3.68. (It will be remembered that the purpose of this study was to test whether or not market research data of this kind followed the familiar Bernoullian distribu-

82 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

OBSERVED PERCENTAGES OF PURCHASE FOR 51 SAMPLES OF 100 INTERVIEWS EACH FOR TWO BRANDS OF DENTIFRICE [3]

BRAND A		BRAND E	
<i>No. of Samples of 100</i>	<i>Percentage who Bought Brand A in Each Sample</i>	<i>No. of Samples of 100</i>	<i>Percentage who Bought Brand E in Each Sample</i>
1	9	1	0
0	10	0	1
2	11	1	2
4	12	3	3
3	13	5	4
4	14	6	5
6	15	9	6
9	16	8	7
6	17	7	8
4	18	5	9
3	19	2	10
1	20	2	11
3	21	2	12
3	22	—	
1	23	51	
0	24		
1	25		
51			
Average of 51 samples of 100	16.39%	Average of 51 samples of 100	6.63%
Actual average for 5165	16.22%	Actual average for 5165	6.56%

tion.) This means that the theoretical variation of one standard deviation would give 16.22 ± 3.68 , or between 19.90 and 12.54 per cent. An examination of the table will show that 35 of the 51 groups of 100, or 68.6 per cent (theoretical percentage is 68.3), fall within this range. If twice the standard deviation, the range is between 23.58 and 8.86 per cent. Fifty of the 51 samples of 100, or 98 per cent (theoretical percentage is 95.4), fall within this range. If we take three standard deviations, the range is from 27.26 to 5.18 per cent and all of the samples of 100 (theoretical percentage is 99.7) fall within this range.

These results, as well as others cited, demonstrate empirically that the binomial distribution and the corresponding probabilities and laws of chance apply to research data of this kind. To the best knowledge of the author this is the first empirical demonstration of this fact and as such represents a definite milestone in the field of marketing research. The resulting certainty of knowing within what latitudes of error to interpret results has made for definitely increased accuracy in the interpretation and understanding of findings based on this kind of market research data.

RELIABILITY AND VALIDITY OF BAROMETERS

In 1936 the Psychological Corporation made five grants for special research projects in connection with the Brand Barometers. Two of these, one for a study of reliability and one for a study of validity, were made to John G. Jenkins, then professor of psychology at Cornell University and in charge of the marketing research work of the Psychological Corporation in Ithaca, Binghamton, and the lower-central area of New York State.

In a regular Brand Barometer survey, 150 interviews were made in B, or upper-middle class, homes. The interviews were then "misplaced." Two days after the original interview a repeat interview was made wherever possible with the same respondent, the explanation being that the interviews had been lost. Exactly 100 repeat interviews were obtained. An item-by-item analysis of these 100 interviews showed the following results:

SHOWING THE PERCENTAGE OF RESPONDENTS NAMING THE SAME BRAND ON REPEAT INTERVIEW [4]

<i>Commodity</i>	<i>Per Cent of Respondents Naming Same Brand</i>
Beer (home)	97
Tomato juice	93
Cigarette (women)	93
Toilet soap (men)	91
Face powder	91
Scouring powder	91
Gasoline	91
Coffee	90
Mouthwash	90
Motor oil	90
Beer (away)	90
Toilet soap (women)	89
Cigarette (men)	89
Tooth paste	88
Face cream	88
Flour	88
Shaving soap	88
Hand lotion	87
Tire	85
Average	90.0
A. D.	1.8

A special analysis indicated that there were no reliable differences in the amount of shifting toward or away from the more widely advertised brands. These high percentages of agreement and the uniformity for different brands certainly indicate a reliability more than adequate for most practical purposes.

In a subsequent study of validity [5], housewives' responses to Brand

84 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

Barometer questions were checked against sales slips in a local store. The following table, based on data from 70 homes, shows the "percentage of agreement" between sales slips and Brand Barometer responses.

SHOWING PERCENTAGE OF AGREEMENT BETWEEN BRANDS REPORTED AS LAST PURCHASED AND SALES-SLIP RECORDS OF ACTUAL LAST PURCHASES

<i>Commodity</i>	<i>Percentage of Agreement</i>
Bread	100
Salad dressing	93
Canned soup	90
Coffee	88
Butter	84
Canned fruit	78
Soap	72
Canned vegetable	71
Crackers	70
Cereal	70
Package dessert	67
Cheese	63
Flour	62
Mean	77.5
A. D.	10.4

It will be observed that the above percentages show "agreement" between individual sales slips and the corresponding Brand Barometer interview. The "percentage of agreement" of the Brand Barometer responses of the entire group with sales-slip figures will usually be higher than this except where there is perfect agreement of individual responses. This is due to the fact that the "non-identical individual responses" tend to distribute themselves in relation to brand popularity. This will be obvious from the following table:

<i>Brand</i>	BRAND BAROMETER RESULTS				<i>Group Responses Identical with Sales Slips</i>
	<i>Sales Slips #</i>	<i>Individual Responses Identical with Sales Slips</i>	<i>Distribution of Nonidentical Individual Responses</i>	<i>Total Group Responses</i>	
A	32	27	5	32	31
B	26	23	3	26	26
C	18	15	2	17	17
D	10	8	3	11	10
E	10	8	2	10	10
F	5	4	0	4	4
	100	85*	15	100	97†

*This would mean 85 "per cent agreement" according to Jenkins' procedure.

†If a "group-to-group" comparison is used, there would be 97 per cent agreement.

The above table is only illustrative, but it indicates that the degree of validity of the group as a whole may be higher than the "percentage of agreement" of individuals. This occurs, of course, because the "nonidentical" responses tend to distribute themselves in the frequency order of the identical responses and of the sales slips.

REFERENCES

1. For more complete accounts of this early work and description of the Brand Barometer see: Link, Henry C., A New Method for Testing Advertising and A Psychological Sales Barometer, *Journal of Applied Psychology*, 18 No. 1 (February, 1934), 1-26; and Link, Henry C., and Lorge, Irving, The Psychological Sales Barometer. *Harvard Business Review*, January, 1935, 193-204.
2. Link, Henry C. How Many Interviews are Necessary for Results of a Certain Accuracy. *Journal of Applied Psychology*, 21 No. 1 (February, 1937), 1-17.
3. *Ibid.* p. 6.
4. Reproduced from Jenkins, John G., Dependability of Psychological Brand Barometers, I. The Problem of Reliability, *Journal of Applied Psychology*, 22 (February, 1938), 1-16.
5. Jenkins, John G., and Corbin, Horace H. Jr., Dependability of Psychological Brand Barometers, II. The Problem of Validity, *Journal of Applied Psychology*, 22 (June, 1938), 252-260.
6. *Ibid.* p. 256.

Chapter VII

PUBLIC RELATIONS RESEARCH

WALLACE H. WULFECK

Federal Advertising Agency, Inc.

THE organized, consciously devised public relations program is a phenomenon of the between-wars era. Public relations programs are the product of "big business" and its necessity for maintaining healthy public attitudes during a period when the climate of public opinion was increasingly unfavorable to large centralized business enterprise. Initially these programs were *ex post facto*. That is to say, they emerged as the direct result of the public's putting industry on the defensive, thereby threatening the sales position of their products in a competitive market or their investment position in the financial market. In general, what usually happened was that a given industry suddenly became aware of a mounting negative, or protest, attitude on the part of the public, or its employees, toward its publicized methods of operation, labor relations, dealer relations, advertising or selling methods.

In the earliest days of developing negative opinion, few industries had either the experience or the personnel capable of constructing a propaganda program effective enough to manipulate successfully the mass psychology of the investing or consuming public. The result was the rise of independent public relations counsels, who were retained by individual industries and frequently by trade associations of industries in related lines to counsel, devise, and execute programs of public education favorable to their clients. Today the evidences of these measures can be found by any discerning reader of newspapers and magazines, or listeners to the radio. No one can question the effectiveness of the job accomplished, but many have questioned the honesty of the tactics employed.

As soon as industry and its public relations counsels started to work, it became strikingly evident that an intelligent policy would dictate the maintenance of a continuous educational program to stave off the possibility of

crises in public attitudes developing. Once this requirement was recognized and accepted, public relations research became inevitable. It became clear that a campaign of public re-education had to begin with accurate knowledge of the present state of the public mind. Moreover, industry wanted to know what was accomplished by their effort and at the same time the public relations counsel was forced to demonstrate to its client some measure of the effectiveness of its program. This meant the development of research devices to measure the continuous trend of stability or swing in public attitude.

Two needs thus provided the foundation for public relations research. First, the need to know what the public mind was at any given moment as a base from which to begin educational operations. Second, the need to measure the relative effectiveness of the programs as they were developed.

Turning now from this brief historical appraisal of the background of public relations research, it would be valuable to attempt a definition of the problem of public relations and the areas enclosed within it. This is necessary if we are to understand the broad field of investigation to be covered by such research.

A definition of the phrase "public relations" is relatively simple, but a description of the areas of influence within which to operate is much more complex. [1] As a matter of fact, there is not a great deal of common agreement on this point. To begin with a definition—*public relations is the process of educating public attitudes in the direction of favorable opinion, or positive approval, of the operations and policies of a given industry through the use of advertising and publicity techniques.* There is plenty of room to quarrel with this definition on the ground that it is too narrow but in essence it covers the basic activities of practically all programs in operation today.

Strong argument can be made for including the adjustment of the actual operating policies of a company to the public expectancy as the basic element in any public relations program, but few companies have accepted this radical notion.

The problem of defining the areas in which a public relations program must work is tied up closely to the conception of what constitutes the "public." In the early days the public was conceived as all those persons not directly related to or identified with the management of a given company. For example, it tended to exclude stockholders and employees. Generally, little attention was paid to the stockholder beyond the mailing of the annual statement of the industry, which was usually obscure and unintelligible. The employees' attitudes, on the other hand, became a problem in "labor relations," with all that the term implies, and was assigned to a separate group of specialists in a department all their own.

Many of the business organizations which have made the greatest progress

88 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

now group these functions into one category and regard them as one and the same problem. They define their areas of public relations work as including:

- (1) The general public
- (2) The consumers of their products
- (3) Their employees
- (4) The stockholders
- (5) Their management
- (6) Government

A program conceived on this broad base recognizes the interrelatedness of all these groups in creating public attitudes and the wide ramifications of industry's responsibility to them.

Research, then, in this field, must contend with the specific attitudes of these various heterogeneous segments of the public and at the same time investigate the ways in which the attitudes of any one segment tends to influence the direction of the attitudes of all of the other segments. At this writing many of the ramifications of these complex interrelationships have never been investigated, partly due to the fact that techniques have not yet been developed to measure reliably the subtle impact of one element upon the other.

USES OF THE QUESTIONNAIRE IN INDUSTRIAL RELATIONS RESEARCH

It was pointed out above that many of the problems in this field have never been subjected to investigation by questionnaire techniques. Nevertheless, the past few years, notably the war years, have seen more frequent application of these methods in limited areas of investigation. Not much imagination or daring has been exhibited in attempting to devise unique applications for digging into the elusive and apparently intangible elements underlying many of the problems. Perhaps these adventures will come only through the pressure of necessity for solving the increasingly complex problems of the future. We can hope they will come before that time.

In outline form the following is a partial list of the ways in which the questionnaire method has been used.

1. Through periodic nation-wide surveys to determine and plot the changes in public attitudes toward specific corporations. These have been trend studies conducted for more than ten years.
2. For occasional and trend studies to measure public confidence in business, labor unions, government.
3. For occasional investigations of public opinion toward a specific industry, a philosophy of government, i.e., the New Deal, a particular labor dispute or a special personality, i.e., John L. Lewis or Sewell L. Avery.

4. For investigating stockholders' attitudes toward the management and policies of the corporations in which they share ownership.
5. For finding out what the worker thinks about his job, his company, his foreman and his labor union.
6. For measuring job security attitudes among workers during the war and projected to the postwar period.
7. In determining public attitudes toward and confidence in advertising.
8. For having employees select from among several alternatives the retirement benefit plan they prefer without their knowledge that the survey is being made among the employees of their company.
9. For estimating the competitive position of a given manufacturer against others independently of the product brands made for sale.
10. For predetermining the attitudes of distributors toward the sale of a product which has never been sold, on the basis of the manufacturers' reputation.
11. Various war agencies of the government have made extensive use of questionnaire studies to estimate public reaction to projected programs and to measure public response to programs already in operation.

With this background in mind we proceed to examine in greater detail several of the more successful ways in which research methods have been applied to measurement in the field of public relations. Since practically all research in this field is private and confidential, little has been published about the methods employed. Not many specific case histories are available for discussion. Therefore, the following material was selected to be illustrative of the techniques employed.

MEASURING TRENDS

Oddly enough, *measuring trends in public attitudes toward a group of large industries* was probably the earliest serious attempt to investigate this field. In the early 1930's the Psychological Corporation established, under the direction of Dr. Henry C. Link, its Brand Barometer. This nation-wide survey was designed to make continuous periodic measures of consumer preferences for certain branded merchandise. However, it was not long before the questionnaires included questions intended to measure the extent of changing attitudes toward a limited group of large corporations.

Interviewing for the Brand Barometer is done in some fifty representative cities each quarter of the year. The sample size of 10,000 is constant from period to period and is carefully laid out to secure a representative cross section of the population. Controls are applied to geographical distribution, city size, age, sex, and economic status of the persons to be interviewed. Since

90 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

it is not a panel operation, a new population is interviewed for each period, assuring randomness of results and avoiding stereotyped responses.

It is not possible to report here the exact wording of the question (which is always the same) nor the names of the companies for whom reports are made. However, we can set up a hypothetical question and a set of imagined results in tabular form simply to indicate the ways in which trends may be measured. Let us assume then that the questions is:

1. How do you feel about the policies of X Company?
 Favorable _____ Unfavorable _____ Indifferent _____
 Don't Know _____ No answer _____.

Obviously this is an extremely broad policy question. The responses to it mask many underlying attitudes toward specific aspects of the companies' policies or program. Responses to such broad questions provide an index of what might be called the "halo" attitude. Depending upon what is wanted, it is possible, of course, to ask questions regarding specific attitudes. For example, we might ask:

2. How do you feel about the labor policies of The X Company?
 Are they fair _____ or unfair _____?
 Don't know _____ No answers _____.

However, it is clear that if we are concerned only with the trend of over-all attitudes toward a group of companies, then the more general type of question provides an adequate basis for measurement provided the question is properly phrased.

The results of question 1 for over-all response of the sample might be tabulated as follows:

	OCTOBER, 1944				JANUARY, 1945			
	<i>Fav.</i>	<i>Unfav.</i>	<i>D.K.</i>	<i>N.A.</i>	<i>Fav.</i>	<i>Unfav.</i>	<i>D.K.</i>	<i>N.A.</i>
Company A	76%	20%	3%	1%	80%	10%	0%	2%
Company B	52	40	8	0	52	40	8	0
Company C	81	10	2	7	59	31	10	0
Company D	46	52	1	1	23	56	14	7

Thus, Company A shows a slight improvement in favorable response. For Company B there is no change in attitude, Company C has lost a significant percentage of public favor, and Company D shows an even more serious position. It must be pointed out that the "Don't Know" and "No Answers" are often as important for interpretation of position as are the unfavorable responses. The plotting of these data on graphic trend lines for long periods shows even more clearly what is happening to each company in relation to the others as well as the individual long-term changes for each participant.

Quite frequently marked shifts in opinion can be directly related to special events reported in the press or on the radio, such as a strike or some remarkable contribution to the war effort. A notable case in point was the Montgomery Ward controversy with the President and the War Labor Board. If it had happened that this company's public attitude history had been measured and plotted, current sampling would undoubtedly record significant changes in the public attitude as a direct result of the publicity given this fight. The company then might find guidance in determining the wisdom of its actions so far as public approval or disapproval might influence its policy.

A further consideration of the tabulation above reveals that detailed breakdowns may provide important material for analysis of attitudes by sex, age, economic status, city, and geographical area. Frequently pockets of negative opinion can be isolated and plans made to re-educate particular areas or groups. Or through the use of special methods applicable to certain classes of the population programs may be developed to shift attitudes into a more favorable frame. Often, in research of this sort there appear wide differences between plant cities and nonplant cities because the interest and information is greater close to home. Therefore, it is always wise to make this type of analysis, since the populations of the two groups (in plant and nonplant cities) are not of the same complexion and the over-all results are often colored by this fact.

There are, no doubt, other techniques being used to plot trends in public attitudes toward industry. The foregoing is one suggestive method for doing the job. It is sufficient for our purpose here to illustrate the application of marketing research methods to the measurement of industrial relations problems.

INDUSTRIAL RELATIONS RESEARCH

Modern leadership of industrial management is more and more adopting a point of view which holds that the venerable paternalism of management never has and cannot now effect a successful rapprochement between their workers and themselves. Nor can government regulation or a politically minded labor leadership effect the necessary relationship of mutual trust and confidence. A profound lack of understanding of the purposes, methods, and aspirations of each other is the root of the trouble between workers and management.

Organized labor has as yet made little effort to find a basis for a working relationship. It is committed to a policy of coercive conflict to achieve its ends rather than co-operative understanding. Except for the method of conflict, the same might have been said of business until recently. Only in the past

few years several groups speaking for management have made serious attempts to get at the underlying attitudes of workers through research. At least management seems to show a dawning recognition of the need for mutual understanding.

Two of these researches utilizing different field interviewing methods will serve as examples of the way in which these techniques may be applied to investigations of industrial relations.

The first of these studies has been going on for several years. As this is being written, plans are completed for continuing the interviewing program well into the future. Under the guidance of Everett R. Smith, director of research for Macfadden Publications, Inc., a program of intensive in-the-home interviewing has probed deeply into the minds of American workers to find out what management must know about their attitudes, needs, and aspirations in order to understand them. This is a sincere attempt to get at the truth. Among others, they sought answers to such questions as "Why is it that the great majority of them [workers] lack confidence in business? Why do they feel that they must look to government? What do they want for themselves and their families in the postwar period? And what can be done about it?" [2]

The answers to these questions did not come as specific responses to formal questions, but were secured through long informal but guided conversations. Considerable time was spent in developing and evaluating the informal interviewing method. The men who did the work became expert in drawing out the workers so that underlying attitudes were spontaneously expressed and the emotional strength of the attitudes revealed. Much of the interviewing was done by Mr. Smith himself.

Interviews were made in the homes of workers adjacent to the plants being investigated. The respondent knew that the study had no connection with their employers. As Mr. Smith points out, these were "discussions where the interviewer has the complete confidence of the worker and his wife, to the degree that they will express freely their personal and private opinions and reactions."

The number of interviews made in any one locality was determined by the pattern of the attitudes revealed. If, for example, 25 or 50 interviews showed a pattern in which common denominators were constantly recurring with little variation in attitude, then this number was regarded as adequate for that area. If, on the other hand, there were wide variations in response, the interviewing was continued until no new facets appeared. Although due regard was paid to the statistical and quantitative requirements for reliability, in essence this is a qualitative enterprise.

The reports of these data are entirely qualitative and interpretive. It is this "reporting" treatment which gives the results their peculiarly down-to-earth

value. These reports are the reality, the stuff of the minds of men and women. As such they bring the problem down to the level of daily living where it actually exists and where understanding must take place.

Because the reports are individual and qualitative, they cannot be adequately summarized here. Suffice it to say that the suspicions toward management methods, the misconceptions, the insecurities, the needs, the hopes, the aspirations, and the expectancies of the workers are outlined in sharp relief. There can be no mistakes about the question of where management stands in the minds of employees. Sometimes the answers are so clear that the corrective program is at once obvious. This method of reporting field research results is a particularly fine example of one way to make data live for the reader. The three reports are generally available and careful study devoted to them will be rewarding to the student who is interested in the applications of field methods and especially the method of presentation and interpretation of results. Many practitioners of opinion research are expert technicians, but few demonstrate genuine competence and imagination in the analysis and interpretation of their findings.

A second survey of worker opinion illustrates the use of the questionnaire method as contrasted with the conversational type reported above. As in the previous case, this study was sponsored by a publication. *Factory Management and Maintenance* employed the Opinion Research Corporation of Princeton, N. J., to interview factory workers across the nation . . . "in the interest of better understanding between management and worker."

A formal questionnaire was developed to be used by the interviewers who stopped industrial workers on the street near a plant. Examples of the types of questions used will be given in the following excerpts from the results.

The sample was of sufficient size to permit a projection of the results to the 17,000,000 wage earners in America's manufacturing industries. Control of the sample can best be illustrated by quoting the report: ". . . great care was taken to allot the interviews in exact proportion to the geographical concentration of industry, type of industry, size of plant, length of employee's service, age group, and sex. For example: 10 per cent of the industrial population resides in New England; 11 per cent of *Factory's* interviews were conducted in New England.

"Another example: 23 per cent of industry is found in cities of 100,000 to 500,000 population; 25 per cent of the interviews were conducted in cities of that class. Still another example: 17 per cent of the industrial population is engaged in textiles—apparel-leather production; and 19 per cent of the interviews were with workers in those industries." In other words, the sample was believed to be "representative of hourly paid and piece rate workers throughout the nation."

Other writers in this volume will have discussed the importance of the adequacy of the sample for reliability of results. We will not repeat here except to point out the essentiality of a true representation of the population to be measured. Less than that leaves the results open to question, especially where projections are to be made against a total population as in the case at hand.

The following are typical questions used in the survey:

Question: "In the first year after the European war ends, what sort of picture do you see in the matter of unemployment—do you expect a lot of unemployment, a medium amount, or not very much?"

Answer:

"A lot"	29%
"A medium amount"	24%
"Not very much"	36%
Undecided	11%

Question: "Who do you think will do the most to solve the job problem after the war—the government, labor leaders, or company heads?"

Answer:

Government	47%
Company heads	24%
Labor leaders	14%
Undecided	15%

Question: "Which way to you think the working man has the best chance of making a good living—by joining a union or by trying to go ahead on his own?"

Answer:

Joining a Union	72%
"Go it alone"	14%
Can't decide	14%

Question: "Do you feel that the head men of your company have the workers' interests at heart; or do you feel that they are not much concerned with what happens to the workers?"

Answer:

"Have workers' interests at heart"	56%
"Are not much concerned"	35%
"Have no opinion"	9%

Question: "Would you say the pay of foremen in your plant is about right, too high, or too low?"

Answer:

About right	62%
Too high	6%
Too low	13%
Have no opinion	19%

These questions suggest the general tone and direction of the problems investigated. They tend to reflect the guided opinion of the wage-earning population toward the critical aspects of their job security and where they place their faith in the solution of the problem. Broken down by types of industry, age, union, membership etc., they provide penetrating insights into the minds of workers and set the basis for understanding on which harmonious relationships can be established.

The two foregoing examples demonstrate specific ways in which opinion research has been applied to the study of industrial public relations. From this material it is an easy step for the imaginative research-minded person to envisage the possibilities inherent in these techniques. Not all the answers can be obtained this way, nor can the problems be solved simply by acquiring knowledge of this sort. But without this kind of information on which to build there can never be hope for achieving mutual understanding and harmony between the public, the employees, and industry. Therein lies the challenge for research.

REFERENCES

1. Garrett, Paul W., Public Relations—Industry's No. 1 Job, *Bulletin American Association of Advertising Agencies*, April, 1938.
2. Smith, Everett R., *The Destiny of Free Enterprise, 1943; The Pattern for Survival, 1943; and What Shall We Do About It? 1944.* New York, Macfadden Publications.

Chapter VIII
CONSUMER RESEARCH IN THE DEVELOPMENT
OF ADVERTISING COPY

ALFRED C. WELCH
Knox Reeves Advertising, Inc.

GENERAL PURPOSES

THE purpose of using consumer research in the development of advertising copy is to contribute whatever is possible toward attaining the objectives the particular advertising is designed to accomplish.

Development of the advertising is only one part of the advertising program. Consumer research is only one cog in the process of developing effective copy. Research to develop copy is separated from other advertising research only for convenience in analysis and explanation. In actual practice, other applications of consumer research—such as copy testing and product analysis—are so intimately tied in with research to develop copy that all often are combined into a single research plan.

Often the principal purpose of an advertising campaign is to get people to buy a particular brand and to keep on buying it. Here the research program faces two questions: (1) What should you tell them to get them to buy and to keep on buying? (2) How should you say it? Emphasis should be put on the "keep on buying" objective to prevent heading toward short-run, high-pressure techniques and to focus major efforts on building up long-run consumer satisfaction.¹

¹The principal purpose of the advertising is not always direct promotion of sales. Sometimes it is merely intended to get people to think favorably of the company. The first question simply is rephrased to fit the objective: (1) What should you tell them to get them to think favorably of the company? (2) How should you say it? Similarly, the questions can be varied to fit any objective; and, of course, it is understood that success in attaining the objective is to develop advertising that will do the job at the lowest cost.

Although other questions must be answered in working out an advertising program, these are not connected with copy development. Where should you say it? is a problem of media

Research in the development of advertising copy is not intended to replace imagination, ingenuity, or skill of the advertising strategist, the copy writer, or the art department. It is intended to supply information to help everyone concerned do a better job. So far, it has been more successful in being a real help in finding out what the advertising should say than in suggesting how it should be said.

SPECIFIC PURPOSES

Suppose you suddenly inherit a company equipped to manufacture a certain product. Your company has its sales force, its dealers, and everything else you need except an advertising campaign. What problems would confront your company and its advertising agency in developing effective advertising? Restrict the problem still further by considering only what you should say in your advertising and how you should say it, forgetting for the moment the other questions with which you would be concerned if you actually were in this position. Solution of each problem would be a specific purpose of the use of research in the development of your advertising copy.

1. First of all, you would want to get acquainted with your customers, both actual and potential. Who are they? How do they use the product? Who influences their decision to buy? Under what conditions do they buy? What do they think about your product and about competing products? The better you know your actual and potential customers the better job you can do in influencing them to buy your product and keep on buying it.

2. Next, you would want to find a basic idea to stress in your advertising: the principal idea or ideas you want to get across to your audience. At this stage the idea may be in quite general form. For example, you might decide that you will stress the quality of your product, its nutritional value, its economy, or that it is especially good for a particular use. Research helps you decide which basic idea provides the best foundation for effective advertising.

3. Usually, putting your basic idea directly into an advertisement in its general form will not be satisfactory. To get highest effectiveness it often is necessary to express the idea in specific, concrete form. The basic idea may merely be the impression you want to create or the appeal to be used. Effective presentation may require a specific copy theme or some other concrete expression of the idea. Research helps you select the most effective concrete expression.

evaluation: radio, or newspapers, or magazines, or outdoor advertising, or other media—and the selection of specific radio networks or stations and particular publications. To whom should you say it? comes under the analysis of market potentials. These questions are interrelated. What you should say depends to some extent on to whom you are saying it; and what you want to say may be important in deciding where you should say it. Discussing only what to say and how to say it is somewhat of an abstraction; in actual practice these questions are not treated as independent problems.

4. Now you know what you are going to say in your advertising. How to say it has been partially answered because you have a specific theme or other concrete expression to convey your basic idea. How you are going to say it still needs to be worked out in more detail, through research. The end products are advertisements or radio commercials now ready for use.

5. Before going into a large-scale advertising campaign, testing on a small scale usually is a sensible precaution. Although research has been used during each step in the development of the copy, the advertising has not yet been tried under actual conditions. Since emphasis at this point should be on further improvement of the advertising, this test can legitimately be included as one of the phases in the development of the advertising itself.

6. After the suggestions of the small-scale tests have been considered, your advertising is launched on a full scale. This might be the end of the story, but often analytic tests are put to work again to spot any possible weaknesses in the advertising as a basis for additional improvements.

Sometimes it is safe to skip one or more of the six steps. The first specific purpose—getting acquainted with your customers—often does not call for a special study because the advertiser or his advertising agency may be well acquainted with the actual and potential customers before a specific campaign is planned. Sometimes one survey will cover several of the specific purposes. The get-acquainted survey may come up with information that not only provides the necessary background but also discovers a specific theme which can be translated into concrete form without further research, thus combining two of the steps and omitting the one between. Finally, owing to the pressure of time or for other reasons some or all of the research steps will be omitted in developing an advertising campaign. Many effective advertising campaigns have been produced without the help of consumer research. On the other hand, a substantial number of advertising attempts without research support have failed to get returns that would justify their expense. It is still too early to say definitely whether advertising developed with the aid of research will be superior; but the growing number of advertisers and advertising agencies using consumer research suggests that many people think it will.

Consumer research *helps* in attaining these specific purposes; it does not develop effective advertising by itself. All it does is provide assistance for the departments responsible for the strategic and creative functions in developing the advertising campaign.

CHARACTERISTICS OF STUDIES IN THIS FIELD

The variety of studies possible in the field of the development of advertising copy arises from the differences in the specific purposes. This variety

also is accounted for by the many available methods. Certain characteristics of methods are especially important both in describing the methods and in selecting one for a particular problem.

1. *Degree of control.* Some studies take conditions as they are and either attempt no control at all or depend on statistical techniques, breakdowns, matching, or other substitutes for actual control of the conditions under which the data are collected. Other studies actually control the conditions.

2. *Method of collecting data.*

(a) *Observation.* Sometimes it is possible to conduct a study merely by observing people, by recording responses mechanically, or by some other method that does not require asking questions. An eye camera measures the amount of time people spend on each part of an advertisement without asking questions.

(b) *Questioning.* Asking questions is probably the most common procedure. In this type of study two characteristics are especially important in understanding the kinds of studies conducted.

(1) *"Tightness."* Surveys can vary in the rigidity or objectivity of the questions asked. At one extreme is the fixed question requiring a direct "Yes" or "No" answer or some other straightforward response such as naming the brand purchased last. All the interviewers ask the same questions in exactly the same way and the answer is simply checked or recorded. At the other extreme is the "open-ended interview," also called the "intensive interview" or the "depth interview." The interviewers are assigned a limited number of questions. After the initial answer to each question the interviewer attempts to get a more complete and more systematic response before going on to the next question. He may ask a whole series of questions to get a full report.

(2) *Degree of self-analysis required.* Questions requiring the respondent to analyze his own behavior often are used in studying buying motives and appeals. The "why question" is a good example. People are asked why they prefer a certain brand or why they bought one brand instead of another. Again in the "reason-why check list" respondents are asked which characteristics (economy, durability, appearance, etc.) influence them most in buying a product. Hand in hand with self-analysis goes accuracy of report. Even if a respondent does accurately analyze his behavior he may not report accurately because certain answers may be more acceptable socially than other answers. At the other extreme are questions requiring practically no self-analysis; for example, questions which ask the brand bought last.

(3) *Treatment of the results.* The distinction between two different ways of handling the responses obtained by the study is especially important in the use of research in developing advertising.

(a) *Simple tabulation.* Often the answers are simply counted or classified in some way and the number falling into each category shown.

(b) *Breakdowns.* In breakdowns the responses are divided into different groups, and the groups are compared on some characteristic not used in dividing the responses into the groups. For example, in a taste study one group consisted of those who thought this particular brand was superior in taste and the other group of those who did not think it was superior in taste. Preferences of the two groups for this brand were compared to find whether people thinking the brand is superior in taste also tend to prefer the brand. The breakdown is usually used to study the relation of one factor to another.²

(4) *Process of reaching conclusions.* Some methods produce a quantitative measure which leads to the conclusion directly. Reasons given for buying a product can be counted and the most frequent one accepted as the most important. If the breakdown is used to study the relation of taste and preference, an index can be worked out to show the importance of taste. Then, if this index is higher than similar indices for other attributes (such as color and nutritional value), taste can be accepted as the most influential attribute.

In contrast, other methods may require considerable interpretation and inference in reaching conclusions. To find the relative importance of utility and appearance in hairbrushes, one might observe how many families had their brushes in sight (appearance), how many had only a utility-type hairbrush, how many had both utility and decorative types, how often hairbrushes are bought as gifts (on the assumption that appearance would be important in a hairbrush bought for a gift), etc. From the results of these observations a conclusion on the relative importance of utility and appearance would be reached through inference, with no attempt to reduce the results to a numerical index.

(5) *Problem-coverage.* Two studies identical in all characteristics just explained could still differ in the number or types of problems they are designed to solve. Selective and analytic tests of the effectiveness of advertising differ in purpose. Selective tests tell which is best; analytic tests what to do next. If five advertisements have been prepared and the problem is to select the one that has the best chance of being effective, a selective test might be used.

In addition to providing a basis for selecting one of the advertisements, analytic tests also provide information that points toward specific strong

²There are other methods of estimating the degree of relationship or association without using breakdowns. However, they will not be explained, since the breakdown is the method most commonly used in this particular field. More complicated breakdowns will not be explained either since the principal purpose of including the breakdown in this discussion is because it provides one way to get at the relative importance of factors *indirectly* without asking the respondent to evaluate the importance of the various factors which influence his preferences and purchases.

points and weaknesses and consequently they suggest ways in which the advertisements might be improved.

The preceding explanation illustrates that, in addition to the various possible purposes, the characteristics of methods must be considered in describing the kinds of studies possible in this field. Even for one specific purpose, there can be almost as many different kinds of studies as there are combinations of these characteristics.

Furthermore, thinking of each kind of study in terms of these characteristics helps avoid the error of putting methods into oversimplified categories. For instance, the open-ended interview is often treated as a single method. Actually, when it is viewed in relation to these characteristics, it is obvious that all open-ended interview studies are not necessarily the same. They may differ in the degree of self-analysis required, or in the method of treating results and reaching conclusions.

The explanations of the specific purposes and the characteristics of studies suggest the wide variety of possible kinds of studies in this field. Obviously a complete outline of possible studies for each specific purpose would be too long for purposes of this chapter and would contain too much repetition. Further description of studies will be limited to three topics: (1) studies that might be used in connection with one specific purpose—selecting the basic idea; (2) a few other problems and procedures that are especially relevant to the development of advertising copy; (3) an actual case of the use of research in developing advertising.

SELECTING THE BASIC IDEA

The objective is to find what you should say in your advertising to get whatever results you want. What you should say is intended to be broad enough to cover impressions created without words as well as verbal selling messages and narrow enough to be cut off at the point at which the idea has not yet been expressed in a form in which it can actually be used in advertisements or radio commercials. The basic idea finally selected might be, for example: creation of the impression that the product is dependable by showing it operating efficiently under the most unfavorable conditions; stressing the nutritional value of the product; or a specific slogan to be featured in the advertising. The basic idea is the principal selling message; but it can be in the form of an idea, rather than a concrete expression to be used word for word in the advertising. Its translation into concrete form is the job of other departments.

Selection of the basic idea is characterized by its purpose. It may be included as part of a get-acquainted survey, which also provides information

for other uses. It may be combined with product analysis; in fact, the two are closely related because the characteristics of the product which are most important in getting consumer acceptance often provide the basis for deciding what to say about the product.

A number of different types of studies have been used in the search for a basic idea. Several of them have been mentioned briefly in outlining certain characteristics of methods.

1. *The why question.* The basic idea could be selected simply by asking people why they buy as they do, counting the answers, and selecting the most frequent one as the basis of the selling message. If more people mentioned operating efficiency than any other attribute as the reason for buying their preferred make of car, the advertising would be based on the idea that the particular make is superior in its operating efficiency. There are obvious weaknesses in this question.

(a) The question is ambiguous. People may interpret it in different ways. Some think of the characteristics of the particular make of car. Others think of such influences as the advertising and the salesman. Still others think of the reasons why they did not buy a more expensive automobile.

(b) It requires the respondent to analyze his own behavior and to report accurately the results of this analysis. Both influence the likelihood of reaching a valid conclusion.

2. *The reason-why check list.* This approach asks the respondent to select one or more important factors influencing his choice from a list of specified factors. Properly prepared, this check list can represent a less ambiguous statement than the why question. For example, attributes can be presented in one list, influences in another, etc. The obvious weakness of most check lists is that they fail to include all of the possibilities and the listed answers are given an unfair advantage. This can be overcome by selecting the items from adequate preliminary studies. Nevertheless, regardless of the completeness of the list and the clarity of the problem it presents to the respondent, this approach still requires an analysis of one's own behavior and an accurate reporting of the results of this analysis.

3. *A series of related questions.* Ambiguity can be reduced to a minimum by asking a series of questions directed toward attributes, another directed toward influences, etc. Reasons for disliking one's previous brand can be separated from reasons for preferring one's present brand, and the further isolation of reasons for not accepting other brands can be effected. In addition the respondent's thinking can be guided systematically so he does not overlook certain factors. All this increases the chances of getting a valid answer; yet the respondent may be required to analyze his own behavior and to report the results of this analysis accurately. It all depends on what questions are

asked. If, for example, the method consists of asking a series of why questions instead of a single why question, it does have the advantages mentioned and still it requires self-analysis. As actually used in advertising research this method often does require some degree of self-analysis. Usually the method's proponents have pointed out that the assumption of the accuracy of self-analysis is involved. Another advantage of this method is the avoiding of "Why?" By a series of "What?" "When?" and "How?" an indirect approach can be worked out which not only minimizes self-analysis but also overcomes the tendency of people to give whatever answers they consider most acceptable. [1]

Whether this technique can be used without requiring self-analysis of behavior and without the risk of socially colored responses depends on the problem and the ingenuity of the researcher. Consequently, no attempt will be made to put this method categorically in or out of the class that has these two possible weaknesses. Furthermore, this method, as described here, includes only the use of a series of questions when the results are tabulated and the conclusion is based directly on the frequency of the various responses obtained. If the conclusion is based upon the relation of these responses to use or preference either through inference or by means of a quantitative index a different principle is involved and this type of study will be discussed later under a different heading.

4. *The open-ended interview.* Since the open-ended interview consists of a series of interlocking questions, many comments made in preceding paragraphs apply to open-ended interview studies. In development of copy, however, most frequent uses of this method have tended so far to follow a fairly definite pattern: they have had as their objective the synthesis of a number of observations into a unified hypothesis through a process that Lazarsfeld has called "broad motivational interpretations." [2] For this reason, examples of the use of the open-ended interview are presented separately.

Although a relatively small proportion of automobile buyers purchase convertibles, Dichter found through both an observational study and open-ended interviews that convertibles "are the symbol of eternal youth, coveted and dreamed about by every car buyer, but never put into reality." [3] This hypothesis led to a greater use of convertibles in Chrysler show rooms and provided a basic idea for Chrysler advertisements.

Another open-ended interview study revealed that when trying on shoes many customers felt embarrassed to be seen in stocking feet; they did not like to have others seated too near them and were especially sensitive to the salesman's attitude. These observations were the basis of the hypothesis that the shoe-buying situation tends to cause a feeling of inferiority. [4]

Another study showed the importance of the "fresh start" feeling in the

buying of new products and in other common situations. This idea was translated into copy for Socony advertisements. [3]

Practical people may feel that often these hypotheses are somewhat far-fetched and that in many cases the practical applications could be made directly from the factual observations without the theorizing. However, this approach actually has produced basic ideas for advertising. Since proponents of this method evidently are willing to submit the results as hypotheses subject to further testing by more objective methods, final judgment on the approach should wait until sufficient testing by more objective methods has been completed.

The open-ended interview can also be used to deal with the more specific evaluation of factors influencing preference; to find, for example, whether appearance, taste, etc., is the most influential factor. Use of this method will be illustrated later.

5. *Objective questions with breakdown analysis.* This type of study has special importance in this particular field. The questions can be direct, yet the indirect method of analysis overcomes the possible weakness resulting from self-analysis and inaccuracy of report when the direct questions require respondents to evaluate the importance of factors influencing preferences or purchases. Suppose we wish to estimate the importance of appearance in getting people to buy a particular electric refrigerator. Owners could be asked to evaluate appearance, efficiency, economy, etc., in influencing their preferences for electric refrigerators. If appearance showed up poorly, we would not select the superiority of the appearance of our particular refrigerator as the basic idea. Suppose most people underestimate the importance of appearance, or that many of those who do recognize it are reluctant to admit that appearance is a strong influence because they feel such an answer would not be socially acceptable. Under these conditions, the true influence of appearance would be underestimated.

Under these same conditions, what would happen if we used relatively objective questions *without* breakdown analysis? We might ask: "What make of refrigerator do you own now? What make did you own before this one? In what ways is this refrigerator better than your previous make? What other refrigerators did you consider before you decided on this one? In what ways did this one seem better than the other ones you considered?" Would this approach get around people's inability to recognize the true potency of appearance and their reluctance to admit that appearance was extremely important? The writer honestly does not know and is withholding judgment until the results of further research are available.

The same problem could be approached by asking objective questions and using a breakdown to get at the importance of appearance. Respondents

might be asked to rate the refrigerator on appearance, as well as on a number of other characteristics. Some measure of preference would be used. Then the responses of those rating a refrigerator high on appearance would be segregated and their preferences compared with the responses of those who did not rate it so high. The tendency of high appearance ratings to be associated with high preference would be the measure of the importance of appearance in promoting preference for a particular refrigerator.

If appearance was influential in promoting preferences, but if many people failed to recognize its importance, this method should reach the correct conclusion anyway. Realization of the importance of appearance is not involved in either the appearance rating or the statement of preference, yet its importance will be reflected in the relation of the appearance ratings to preferences. Any reluctance to admit the influence of appearance would not affect the results because the respondent merely expresses an opinion on how the refrigerator rates in appearance; he does not report whether appearance was important in influencing his preference.

One study showed that 20 per cent of respondents thought the cost of fuel gas too high and 20 per cent were afraid it was dangerous. [5] A second study was conducted and a breakdown analysis made. Gas was again rated unfavorably on cost by about 20 per cent and about the same percentage objected to its being dangerous. When a breakdown was made, preference for gas and use of gas were about the same among those who rated it as too expensive and those who did not. Similar analysis showed that danger was a much more influential factor. Those who rated it unfavorably on safety had a much lower preference for gas and used only about half as much as the people who did not consider it dangerous. In terms of an index [6] safety appeal was about seven times as potent as economy appeal in getting people to use fuels other than gas. [5]

This method has the advantage of eliminating difficulties of other methods requiring self-analysis and which may be subject to inaccuracy of report. However, it is by no means perfect. It requires a larger number of interviews because the data are divided into groups. The tendency of people to rate their preferred brands high on most characteristics (the "halo effect") complicates the interpretation. At a particular time, the principal brands may be very similar on some attributes and differ widely on other characteristics. For example, the leading brands may sell at the same price. If people realize this, the influence of price may show up as being very low, whereas its real influence might be great in the sense that a good lower priced brand might attract large numbers of customers. This may cause some difficulty in interpretation and in some cases it limits the usefulness of this method.

6. *Observational studies.* Sometimes it is possible to select a basic idea by merely observing people and not asking any questions.

In one study people were observed in front of a table display of bars of soap. Most people slid their fingers along the surface of the soap, then smelled it. Dichter suggests that this discovery of the personal relation between soap and consumer could be translated into the copy theme: "A cake of soap for every family member." [7]

Dichter also discovered the high interest value of convertibles by counting the people who stopped to look at automobiles in showrooms with convertibles and in showrooms with no convertibles. This, with other studies, led to the use of convertibles in Chrysler showrooms and advertisements. [8]

Experimental studies often depend mainly on observation rather than questioning. However, such studies will be discussed under different headings.

Observational studies possess objectivity. Their less frequent use occurs mainly because situations which make it possible to solve the problem by observation alone often are difficult to find in this field.

7. *Experimental studies.* A study of this type will be included in the description of an actual case of the development of advertising copy. This approach to the problem has not been used much. This is partly due to the difficulty and expense of setting up such an experiment under conditions paralleling actual conditions of advertising. Nevertheless, this method does appear to offer possibilities for a number of problems. Research ingenuity may find some way of making greater use of it in the future.

What does all this mean? The principal purpose of discussing certain methods suitable to search for a basic idea was to illustrate the kinds of studies possible. A few of the problems involved in the use of each method were mentioned. For example, the question whether a particular approach requires self-analysis came up frequently. It is difficult to tell whether some questions really require the respondent to engage in considerable self-analysis. It may be more difficult to decide if his report will be distorted by a reluctance to give a socially unacceptable answer. Finally, there is little acceptable evidence on either side of the question of whether people can accurately analyze their own behavior to evaluate the relative importance of the factors influencing their preference.

Authorities fail to agree. Link believes that if people cannot give the reasons for their behavior almost immediately, probing for additional reasons usually results in getting rationalizations and false answers. [9] Smith and Suchman are more optimistic about the ability of people to report the factors influencing their purchases, although they realize that the issue is by no means settled. [10] They used a series of questions to find what influence the Boake Carter news broadcasts had exerted on people who bought Philco

radios. Influence of the program was greater after a change from standard to daylight time. Studies using objective questions on listenership and Philco ownership and using a breakdown analysis also showed that relationship between listening and owning a Philco was higher after the program had changed time; to this extent the objective, indirect, breakdown approach corroborated the conclusions of the studies which used questions requiring some degree of self-analysis. Although this study dealt with influence of a radio program rather than selection of a basic idea, it is relevant because it provides evidence on the issue of the ability of people to analyze their own behavior and to report the results accurately.

Lazarsfeld has reviewed three studies dealing with this problem. [11] In these, the reasons given held up well when checked against the results of the breakdown analysis.

Although the tendency of people to avoid reporting socially unacceptable reasons is generally recognized, there is little objective evidence even on this question. Of course, if a self-analysis method agrees with a method that does not require self-analysis, it is logical to conclude both that the respondents were able to analyze their behavior and that they reported accurately. Separation of the two different possible sources of error is seldom feasible if the self-analysis method fails to check with another method.

An example from public opinion studies applies here. If one group of respondents report their answers to interviewers while the people in a comparable group report on a secret ballot, differences probably arise because people are less likely to hold back what they consider socially unacceptable answers if they can report secretly. Turnbull found differences between secret and nonsecret ballots which, when the respondents' answers might affect their prestige, were large enough to question the validity of nonsecret reports. [12] On a question about England's carrying her share of the war burden, the nonsecret interview showed a higher percentage expressing the socially acceptable opinion.

In discussing the methods of selecting a basic idea, self-analysis and inclusion of questions subject to socially colored responses were termed "possible" weaknesses, because the validity of self-analysis methods has not been definitely settled. What should be done when one wants to use consumer surveys in his search for a basic idea?

Common sense suggests avoiding the use of methods requiring self-analysis whenever a method requiring less self-analysis can be used. However, when only a method requiring self-analysis seems to be suited to a particular problem, careful use of such a method should yield more valid conclusions than not using research at all. In this case, questions should be selected and phrased to hold self-analysis to a minimum, every effort should

be directed toward getting answers which reflect what the respondents really believe instead of what they think is socially acceptable, and interpretation should make whatever allowances are necessary for possible distortion resulting from errors in self-analysis and tendencies to avoid reporting socially unacceptable answers.

SELECTED PROBLEMS AND PROCEDURES

Many of the more important problems and procedures have been covered in explaining the characteristics of methods and the selection of the basic idea. Still others will be included in explaining an actual case of the use of research in developing an advertising campaign. Two remaining topics are important enough to deserve special mention: analytic tests, which will be covered in this section; radio commercials, which are discussed in another chapter.

1. *Analytic tests.* The difference between analytic tests and selective tests was explained previously. Further description of selective tests are found in the chapter on copy testing.³

Analytic tests do more than select advertisements. They attempt to discover specific weaknesses and thus provide some basis of further improvement, often by measuring different aspects of the effects of the advertising through the use of several different tests. If the problem is confined to getting an indication of the general effectiveness of the advertisements, use of several tests often leads to about the same conclusion as a single test because an advertisement which rates high on one test usually tends to earn a high rating on the other tests, too. [13] The extent to which this applies, of course, depends upon the particular tests used. If the object of the study is to get clues for improving the advertising, in addition to estimating its general effectiveness, use of additional tests often proves worth while (but only if the correct combination of tests is used). For example, if the respondents merely judge the attention value, etc., results may be of little help in pointing toward specific weaknesses. The advertisement rated highest on attention value may also rate highest on the other characteristics. Similarly, each other advertisement may rank about the same on other characteristics as it does on attention value.

However, if tests are used which measure different aspects and if all of them are not based on judgments of the respondents, the chances of getting

³ Special mention of analytic tests in connection with the development of advertising copy is not intended to imply that they are more important than selective tests. Copy can be improved by a process of gradually eliminating the least effective advertisements, using only selective tests. This use of copy testing is extremely important in the development of advertising copy. It is not explained here simply to avoid duplication with the chapter on copy testing.

some indication of specific weaknesses are fairly good. Comparison of the results of analytic tests for two fountain pen advertisements illustrates how analytic tests are used.

Two advertisements for different brands were tested by exposing people to the advertisements and measuring the increase in brand preference by means of a Brand Preference Scale. This scale expresses the degree of preference for a brand in terms of a numerical index. [14] The Brand E advertisement increased preference significantly, the Brand S advertisement did not. *If the only important problem is to decide which of these two advertisements is more effective, further tests would be unnecessary. If the problem is to discover what is wrong with the Brand S advertisement, additional tests are helpful.*

In this study three additional tests were used, [15] and the specific weakness of the Brand S advertisement was discovered by studying the interrelations among the four tests by use of breakdown analyses. Of the seven comparisons made, two provided the key to its weakness.

In one comparison preference of the people who believed the advertising claim was compared with preference of those who did not. Belief in the claim was measured indirectly, not by bluntly asking people whether they believed the claim. For both advertisements, people who believed the claim had a significantly higher preference for the brand making the claim than the people who did not believe it. The results are shown in Table 1.

TABLE 1
BRAND PREFERENCE AMONG PEOPLE WHO BELIEVED AN ADVERTISING CLAIM AND THOSE WHO DID NOT

	<i>Brand E</i>	<i>Brand S</i>
Brand Preference of People Believing the Claim	66*	91*
Brand Preference of People Not Believing the Claim	39*	71*
Difference	27	20

*This is a Brand Preference Index, not the percentage preferring the brand. For purposes of this explanation, it is necessary to mention merely that a high index shows high preference for the particular brand. Anyone interested in further details can find a more complete description in the reference previously mentioned.

Evidently, then, if Brand S can get people to believe its claim, it will increase the preference of these people for this brand. The weakness does not lie in stressing a point that is of little importance in consumers' preferences for fountain pens. This finding in itself suggests the answer: people who believe the claim do have a higher preference, yet the test of general effectiveness showed that this was a weak advertisement. Perhaps the trouble is that the advertisement fails to get people to believe the claim.

Another comparison confirms this conclusion. For the Brand E advertise-

110 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

ment, the respondents who could give the correct answer when asked which brand used this theme were separated from the respondents who could not name the correct brand, and a much higher percentage of the former group believed the claim. In direct contrast, the same analysis for Brand S showed that there was little difference between the two groups in the percentage believing the claim, as shown in Table 2.

TABLE 2
BELIEF IN ADVERTISING CLAIM AMONG PEOPLE WHO WERE FAMILIAR
WITH THE THEME AND THOSE WHO WERE NOT*

	<i>Brand E</i> Per Cent	<i>Brand S</i> Per Cent
% Believing the Claim Among People Familiar With the Theme†	42	31
% Believing the Claim Among People Not Familiar With the Theme†	3	26
Difference	39	5

* Actually, chi-square was used to test the association between theme familiarity and theme credence. For Brand E, P was less than .01 and for Brand S it was greater than .6. Presenting the results in terms of percentages is used here for the sake of simplicity.

† People familiar with the theme are those who can name the correct brand when asked which brand of the product uses the particular theme.

Presumably, the people who can identify the brand using a particular theme are those on whom advertising has made a substantial impression. If their belief in the claim is not considerably stronger than the belief of people unfamiliar with the theme, it is likely that the advertising is weak in getting consumers to believe its claim. Evidently, this was what was wrong with the Brand S advertising. Knowing this, one could attempt to make the claim more convincing.

2. *Radio commercials.* The development of radio commercials, as distinguished from the development of publication advertising, has not received special mention in this chapter. Basically the problem of what to say and how to say it is the same for radio as for any other advertising medium. Further discussion of radio commercials can be found in Chapter 10.

AN ACTUAL CASE⁴

Problem. Bez was an established brand of butter.⁵ Its makers had developed a new process of enriching their product. One variation of this

⁴The studies described in this section were conducted by Knox Reeves Advertising, Inc., under the direction of Dr. A. R. Root, director of media and research, who suggested and designed most of them. However, the description and interpretation presented here, including any errors, are the responsibility of the writer.

⁵Bez is a fictitious name, and actually the product was not butter.

process produced a butter differing in color from the regular Bez, while another variation produced a butter of the usual color. The problem was reduced to three basic questions: (1) Should nutritional value be selected as the basic idea or should the appeals previously used by Bez be continued with nutritional value used merely for supplementary support? (2) If nutritional value is selected as the basic idea, will people accept the enriched product if it looks and tastes just like regular Bez (in which case one process would be used and the butter called Enriched Bez) or will the appeal be stronger if the color is changed to offer tangible evidence of a physical difference and to suggest higher nutritional value (in which case the other process would be used and the butter called Dark Bez)? (3) In any case, in what concrete form should the idea be expressed?

Research plan. In terms of the six specific purposes previously explained, the first step was merely skipped. The necessary background had been built up already through a continuous research program for the regular Bez. The plan started with the problem of evaluating nutritional value as a possible basic idea. At the time, there were no enriched butters in common use. On the assumption that answers given by people who did not have a background of actual experience might not be valid, a method was selected which would provide the necessary experience. This decision led to the use of an experimental study. To get more information and to take advantage of this opportunity to check questioning methods against the results of an experimental study, the plan included an adaptation of the open-ended interview and several objective questions to be asked at the conclusion of the experimental study.

Selecting the Basic Idea. 1. Method. Two comparable representative groups of families were selected and each family was supplied regularly with three butters: Regular Bez, Enriched Bez, and Dark Bez. The butters were prepared under carefully controlled conditions and packaged in containers identified by symbols known to be equal in their preference value.⁶

Families in both groups were supplied with the three butters during a conditioning period. They were given only one each time and the order was controlled so that each butter was tried first, second, and third the same number of times and each one preceded and followed each of the other two the same number of times. This was done to prevent giving an advantage to any butter due to its position in the series or as the result of contrast effects.

The free-choice period followed the conditioning period. The families were supplied with butter regularly, each family receiving whichever butter it wanted each time. The choices during this period provided the indication of consumer acceptance.

⁶ For example, numbers might be unsatisfactory because 1 suggests top preference and letters were not used because A might suggest a higher quality than B.

112 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

For the experimental group cards were attached to each package of butter. The card for Regular Bez (identified in the test by a symbol, not its trade name) explained its good flavor, quality, and food value. The cards for the other two butters explained these same characteristics, but in addition they stressed the nutritional value resulting from the new process and what this nutritional value meant in terms of promoting good health. The card for each butter accompanied it at all times and was displayed on the table when the butter was being used. Consequently, the family members had ample opportunity to become familiar with the appeals presented on the cards.

For the control group no cards were used.

The study was conducted under the name of a research organization so that the families could not identify the product as Bez. A control and an experimental family were never selected in the same immediate neighborhood, since someone in the experimental family might tell a member of the control family that two of the butters had been enriched by the special process. For the same reason, each investigator was confined to only one of the two groups.

At the conclusion of the free-choice period personal interviews were conducted. An adaptation of the open-ended interview started out by asking the housewife the differences between each butter and the other two. Then the interviewer talked informally with the respondent in an attempt to get an indication of her specific attitudes toward each butter with special emphasis on the reasons for liking the preferred butter. Leading questions were not permitted. Objective questions were asked on which butter was best on a number of specific characteristics.

2. *Results.* The principal purpose of the study, of course, was to find whether the nutritional appeal would be effective and whether a change in color would intensify its strength. An incidental purpose was the comparison of the experimental results—the preference shown during the free-choice period—with the results obtained by questioning. These two different aspects will be discussed separately.

(a) *Experimental results.* The effect of exposure to the appeals on the preferences of the families is shown in Table 3.

TABLE 3
RELATIVE PREFERENCES FOR THE THREE BUTTERS IN THE
CONTROL AND EXPERIMENTAL GROUPS

Type of Butter	Relative Preference (Regular Bez = 100)	
	Control Group	Experimental Group
Regular Bez	100	100
Enriched Bez	104	198
Dark Bez	61	137

For every time one pound of Regular Bez was chosen by the experimental group families, almost two pounds of Enriched Bez were selected; but in the control group these two butters were selected with approximately equal frequency. Evidently the cards explaining the higher nutritional value of Enriched Bez did have a substantial effect.

Dark Bez also was more popular in the experimental group than in the control group, presumably as the result of the nutritional appeal, but obviously some other factor was unfavorable. Its color was the only known reason for this relatively poor showing, and the interviews with the housewives indicated clearly that its color was a serious disadvantage.

The conclusion was direct and simple: the nutritional appeal had a substantial effect and, although it was strong enough to counteract the weakness of the color of Dark Bez, its use for Enriched Bez plus the additional favorable factor of the color of Enriched Bez provided the winning combination.

(b) *Results of questioning.* Both in the adaptation of the open-ended interview and in spontaneous comments made during the experimental study, taste was mentioned much more frequently than nutritional value as a reason for preferring a particular butter in preference to the others. If the decision had been made on this basis alone, taste would have been selected as the basic idea. Yet we know from the results of the experimental study that the cards presenting the nutritional appeal really accounted for the greater preference for Enriched Bez and Dark Bez. Maybe taste was considered a more socially acceptable reason or perhaps the respondents merely did not realize what had influenced their preferences.

At the end of the experimental study the respondents also were asked which of the three butters was best on each of a number of specified attributes; for example: Which of the three butters tasted best? The results for three of the characteristics are given in Table 4.

TABLE 4
RATINGS OF THE THREE BUTTERS
ON THREE CHARACTERISTICS IN THE EXPERIMENTAL GROUP

Type of Butter	% Naming Each Butter as Best on Each Characteristic		
	Taste	Nutrition	Color
Enriched Bez	44	26	62
Dark Bez	39	70	4
Regular Bez	17	4	34
Total	100	100	100

Judging from the results in this table and knowing the preferences for the three butters in the experimental group, what conclusion would one reach in selecting a basic idea? Of the three characteristics, color evidently is by far

the most important in explaining the superiority in preference of Enriched Bez over Dark Bez. (As far as is known from all the information including the reasons given, the physical appearance of the butters, and the factors actually introduced into the situation in the experimental study, this conclusion is correct.) In explaining the superiority of Dark Bez over Regular Bez, nutritional value is the most important of the three factors by a wide margin. It evidently was strong enough to outweigh the unfavorable color of Dark Bez. (This conclusion also checks with the experimental results.) Taste might be given some weight, too, but surely nutritional value would be accepted as the most important factor. In all probability, any experienced analyst would select nutritional value as the most likely explanation of the superiority of Enriched Bez over Regular Bez. (This is correct, too, as shown by the experimental study.)

There is some chance, however, that taste and color would be accepted as fairly important in accounting for the superiority of Enriched Bez over Regular Bez, and perhaps part of the superiority of Dark Bez over Regular Bez might have been attributed to taste. This, of course, is incorrect as shown by the experimental study.⁷

The point of this example is not that the experimental approach is always best or that a self-analysis method will usually lead to false conclusions (both the adaptation of the open-ended interview and the spontaneous comments involved at least some degree of self-analysis). For another product or under different circumstances, an open-ended interview might bring out the importance of nutrition. Furthermore, a more searching open-ended interview might have obtained different results. The adaptation used in this study was less intensive than the usual open-ended interview. Getting ratings on attributes and relating them to the preferences as was done in this study (or by using breakdown analyses as previously explained) might well have led to the correct conclusion.

For this particular problem under these specific conditions the experimental approach produced results that could be interpreted more easily and with greater confidence of being right. Especially in evaluating such intangibles as taste and nutritional value, it is a good idea to remember that the reasons people give may not truly reflect the influence of the appeal and that even an indirect approach based on objective questions may be subject to some error unless the results are skillfully interpreted.

Concrete expression of the basic idea. Now it has been decided that the

⁷ What probably happened was merely that the respondents who preferred Enriched Bez as the result of the nutritional appeal tended to attribute other favorable characteristics to the butter they preferred, just as people tend to rate people high on most characteristics if they like them.

nutritional value of Enriched Bez will be the basic idea. Of course, the advertising campaign could consist merely of shouting: "Bez is Best for Better Health." However, there are a variety of ways to express the basic idea of nutritional value in specific, concrete form; and it is reasonable to assume that a better way can be found than merely repeating the idea as a blunt statement. For this reason, special studies were conducted to select one of several ways of expressing the idea in concrete form. Only the outdoor advertising campaign will be used as an example.

1. *Method.* Several poster layouts, each presenting the basic idea in a different way, were tested. The respondents were exposed to the layouts under conditions that were as similar as possible to the actual situation in which people see outdoor advertising.⁶ Order of presentation was controlled so that no poster had an unfair advantage of being seen before or after the others. Respondents expressed degree of interest in each poster; they reported whatever they could remember; and finally, they explained the principal selling message of each poster as they understood it.

2. *Results.* These ratings were used to select the best posters and also to suggest ways to improve them. In one case, for example, a poster rated high on two of the tests but not so high on comprehension. The poster was based on two characters, and one of them was speaking to the other. However, it was not clear which of the two was making the statement. A simple change cleared up this difficulty. In another case, the tests showed that one of the posters was lower in its recall value than in other respects. It turned out that the situation represented in this poster was familiar to only a limited part of the audience. The use of a more generally familiar situation improved the effectiveness. As a check on the possibility that even the best of these posters was not very good, one of them was tested against several award-winning posters of other products. The Enriched Bez poster came out second on this test, only one point behind the poster that rated highest—a negligible difference.

Research on the more detailed features. For the outdoor advertising no further studies to help work out the detailed features of the posters were needed. For magazine and newspaper advertisements, however, two problems came up for which research was appropriate. One question concerned the addition of a comic strip to the advertisement. The other problem involved the choice between a conventional layout and a different type of layout that had been suggested.

1. *Method.* For both problems the Purdue Eye Camera was used. [16]

⁶ For example: the posters were presented to give the same effect as if the person had been riding by an actual 24-sheet poster at a distance of 100 feet at 20 miles an hour and could see the poster only as long as it took him to cover approximately 100 feet at this rate.

This camera takes pictures of the reader's eyes and part of his face as he leafs through a magazine or newspaper. Concealed in a box above the reader's head, it picks up the image from the reflection on a partially silvered mirror through which the reader is looking. The reader does not realize he is being photographed. The record is a series of pictures (on motion-picture film) showing the eyes, part of the face, and enough of the page to identify it. The film is projected on a screen, and the particular part of the page at which the reader was looking when each picture was taken is judged from the position of his eyes and face—a judgment requiring considerable training and experience. The record shows the order in which each person looks at each area and provides an estimate of how long he looked at each area (since the pictures are taken every one-third of a second or at some other regular interval selected by the investigator).

In the first study (the comic strip problem), different groups of respondents were exposed to the different Enriched Bez layouts and to another set of advertisements and comic strips. Thus each group saw the same material with the exception of the one Enriched Bez layout. The length of time each person spent on each page was left up to him.

To compare the conventional type of layout with the other type that had been suggested, two advertisements for another product were found that had identical elements. In one the conventional type of layout was used. In the other the elements had been rearranged to form the type of layout suggested for the Enriched Bez advertisements. These two advertisements were studied instead of preparing special Enriched Bez advertisements or layouts. It was realized that there was a possibility that the results could not be applied directly to the Enriched Bez problem; but it was felt a comparison of two actual advertisements, each in the magazine in which it appeared, would at least shed some light on the problem.

2. *Results.* A comparison of the time people spent on the different areas of the layout with the comic strip and the one without a comic strip showed that whatever the comic strip added in interest value was more than offset by its drawing readers away from the principal selling message. A test of what the readers had remembered and their understanding of the principal points verified this conclusion. Comparison of the conventional layout with the different type that had been suggested showed only slight differences.

Later tests. Before the advertising was used nationally, two different studies were conducted as a small scale test of outdoor advertising under actual conditions. To avoid overlapping with the chapter on copy testing, these studies will merely be mentioned: one tested the extent to which consumers had seen and remembered the posters; and the other one tested the sales effectiveness in terms of consumer purchases. The problem of testing advertising after it

has been used nationally will not be discussed at this point because it can be explained more appropriately under the topic of copy testing; and furthermore, the use of analytic tests, which might be used if emphasis on improvement is desired, has already been covered in this chapter.

The problem of developing advertising copy is not usually attacked with a complete research program of all six steps. For a new product or when a complete research program is needed for some special reason, all six of the specific purposes may be included in the research plan; but usually the problem does not require all of the steps.

Research never develops copy. Research merely provides information to other departments of an advertising agency in the development of the advertising campaign.

The question of how valuable this information is will be left to the reader. In reaching a decision on this question, be sure to include as one of the considerations: How else could one find a better solution to the problems suggested by each of the specific purposes?

REFERENCES

1. Jenkins, J. G. The questionnaire as a research instrument. *Transactions of the New York Academy of Sciences*, 1940, Series II, 2, No. 5, 20-22.
2. Lazarsfeld, P. F. The controversy over detailed interviews—an offer for negotiation. *Public Opinion Quarterly*, 1944, 8, 38-60.
3. Dichter, E. Raw material for the copy writer's imagination through modern psychology. *Printers' Ink*, March 5, 1943, 63-68.
4. Lazarsfeld, *op. cit.*
5. Houser, J. D. Measurement of the vital products of business. *Journal of Marketing*, 1938, 2, 186-189.
6. Strong, E. K. Jr. *Psychological Aspects of Business*. New York: McGraw-Hill, 1938, 470-471.
7. Dichter, *op. cit.*
8. *Ibid.*
9. Link, H. C. An experiment in depth interviewing on the issue of internationalism vs. isolationism. *Public Opinion Quarterly*, 1943, 7, 267-279.
10. Smith, E., and Suchman, E. Do people know why they buy? *Journal of Applied Psychology*, 1940, 24, 673-684.
11. Lazarsfeld, P. F. The statistical analysis of reasons as research operation. *Sociometry*, 1942, 5, 29-47.
12. Cantril, H., *Gauging Public Opinion*. Princeton: Princeton University Press, 1944, 77-82.
13. Advertising Research Foundation. *Copy Testing*. New York: Ronald Press, 1939, 10.
14. Welch, A. C. A test of competitive advertising. *Journal of Applied Psychology*, 1941, 25, 113-123.
15. Welch, A. C. An analytic system of testing competitive advertising. *Journal of Applied Psychology*, 1941, 25, 176-190.
16. Karlake, J. S. The Purdue Eye-Camera: a practical apparatus for studying the attention value of advertisements. *Journal of Applied Psychology*, 1940, 24, 417-440.

Chapter IX

COPY TESTING

ALBERT D. FREIBERG
The Psychological Corporation

IN SO FAR as the ultimate objective of most advertising has been to sell merchandise, the final and most valid test of most advertising is in terms of sales. This statement seems a simple truth, but the application of sales tests to advertising is complicated, expensive, slow—anything but simple. Furthermore, there are many times when sales are the *deferred* objective and where the *immediate* objective of a particular advertising campaign is the correcting of some false impression, the establishment of a trade mark, the creation of goodwill, the keeping alive of a brand name, etc. This emphasis on functions of advertising other than sales has been especially pronounced during the past few years when there has been a limited amount of merchandise to sell.

INTRODUCTION

The field of copy-testing research is large and complex. During the past fifteen or twenty years this field has been increasingly active. Numerous experiments and tests have been carried out, and there is every indication that continued interest and activity will provide at least some of the missing information in this stimulating and important field of advertising. For there are still considerable confusion, lack of knowledge, and disagreement among psychologists in regard to the uses, merits, and also technical details of the various techniques.

Copy-testing is an attempt to give a scientific answer to age-old questions: How many people will see my ad? Or hear my commercial? How many will read it or listen to all of it? How many of them will buy my product or think better of my company? etc. This chapter will describe briefly the principal

copy testing methods used today in testing magazine and newspaper advertising under the following headings:

1. *Recognition and identification tests.* Results are reported as the number of people who noticed an advertisement, the headline or illustration, or the number who read some or all of an advertisement.

2. *Recall tests.* Here the scores are the number of people who remember having seen the advertising of a particular product, who remember a slogan, etc.

3. *Controlled opinion tests.* Here the score is in terms of the advertisement "most interesting," "liked best," etc.

4. *Inquiry and coupon tests.* The criterion is the number of inquiries, the cost per inquiry, or both.

5. *Sales tests.* The criterion is usually the cost per sale, the rate of sales, etc.

Differences in what tests measure. It is immediately obvious from the above that these different tests measure different things. Obviously then, a group of advertisements or an advertising campaign might not do equally well on all the tests. Here is, for example, a table showing rank orders of recognition, recall, and controlled opinion tests of eight advertising themes. [1]

Theme	3 Recall Tests			Controlled Opinion Test	Recognition Test
	I	II	III	IV	V
A	1	3	1	1	5
B	3	1	5	2	6
C	5	2	2	3	2
D	2	4	4	5	1
E	6	5	3	4	4
F	4	6	7	8	3
G	7	8	8	6	7
H	8	7	6	7	8

From the above results it can be seen that the three recall tests rank the themes much as does the opinion tests. The recall tests rank the first three advertisements at the top and the lowest three at the bottom in 15 out of 18 rankings. These results do not correspond with the results of the recognition test. The actual correlations between the results of these tests, as computed by the rank difference method are:

	3 Recall Tests			Opinion Test	Recognition Test
	I	II	III	IV	V
I					
II		.59	.40	.43	.40
III			.60	.76	.20
IV				.70	.24
					-.21

The ranking in tests I, II, and III are based on 200 cases; those in tests IV and V on 300 cases. Inasmuch as there are only 8 ranks in each test the probable errors of the above correlations are high. The significance of these results does not, however, depend on any one correlation but rather upon the consistent trend showing a low correlation between the recognition test and the recall and opinion tests as contrasted to the relatively high correlations among the recall tests and between the recall tests and the opinion test. It is obvious that the first four tests are measuring some things in common, but not at all whatever the recognition test may be measuring.

Although there are many instances where results of different tests correlate highly, this is by no means the usual situation. The correlations range from high positive down to no correlation, and even over to negative correlations. The reason for this is simply that the tests measure different things, and that since advertisements are not uniformly "good" or "bad" in all respects (initial attention, sustained attention, interest, memorial value, immediate action power, etc.) the correlations between tests vary downward from the maximum possible +1.00.

What tests to use in a specific situation. Obviously the tests used should be selected according to the objectives of the advertising. If the purpose of a Buick advertisement is to tell and convince the motoring public that Buick is the best car for the money, then a modified Controlled Opinion Test (e.g., What was the story this advertisement told?) and a test of credibility are indicated. If the purpose of an advertisement is to show a picture of a new product or package, then a readership test or a recall test may be used. However, the purposes of advertisements and advertising campaigns are not often simple and easy to define, and consequently the selection, use and interpretation of tests and test results are much more difficult.

RECOGNITION AND IDENTIFICATION TESTS

Two continuing services—Starch and Grant—offer such studies on a continuing basis for magazines and newspapers.

These tests consist essentially in leafing through a newspaper or magazine with an admitted reader and recording in detail what the respondent says he has seen or read. In the identification test the brand name and other identifying features are blocked out. After the respondent has reported seeing an advertisement, he is usually asked to identify the brand or product and/or the advertiser. Results are usually reported in terms of the percentage of readers who have (1) seen; (2) read some; and (3) read most of an advertisement. In addition, percentages of seen or read are given for different parts of the advertisement. In the identification test percentages of people who can name the sponsor or brand are also given.

Uses. During recent years recognition and identification tests of magazine and newspaper editorial and advertising content have become increasingly popular. Among the factors evaluated on the basis of test results are: size; color vs. black and white; comic vs. display advertising; position in magazine; length of headline; type of illustration; type of appeal; comparison of media—one magazine vs. another, one newspaper vs. another, magazines vs. newspapers, etc.

Reliability. Correlations of reliability for these techniques range around +.90.

Validity. Respondents "recognize" or "identify" advertisements and articles which they have never seen before. Some of this is deliberate falsification. Some erroneous recognition is due to other advertisements, displays, or signs seen which have features common or similar to the advertisement being tested. Other factors influencing test scores are: the length of time between the original seeing of the advertisement and the time of testing; the number of items tested; and the position of a particular item in the series. The principal characteristic of these errors, and one which cannot be avoided as these tests are usually handled, is that the errors vary in magnitude from advertisement to advertisement. For example, the error for Advertisement A may be 5 per cent; and for B, 25 per cent. Therefore, individual scores cannot be "corrected" as these tests are usually handled.

Lucas [2] pointed out the variable characteristics of these errors, their nature and size, and listed the factors which produce them. He also developed a "Controlled Recognition Test." By his method, advertisements are tested before they appear in a publication to determine the inherent error characteristic of each advertisement. The advertisements are tested again after they have appeared. This makes possible a correction of the after test score for each advertisement in terms of its before, or error test score.¹ This is the only accurate way of approximating the "true" recognition test score for a particular advertisement.

A study by Lucas and Murphy [3] is well worth examining at this point. By special permission, copies of the *Saturday Evening Post* were secured before the general release on Tuesday. Approximately fifty advertisements were taken from this issue and mixed with fifty other advertisements of the same size already published. On each Monday of successive weeks preceding

¹ The formula is:

$$\% \text{ Readership} = \frac{\% \text{ YES on post-test} - \% \text{ YES on pre-test}}{100 - \% \text{ YES on pre-test}} \times 100$$

For example, 20% said YES they saw a Colgate advertisement in the pre-test and 40% said YES they saw it in the post-test. The score, instead of being 40% would be:

$$\frac{40 - 20}{100 - 20} \times 100 \text{ or } 25\%$$

122 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

the release of the *Post* on Tuesday, interviewers questioned men and women *Post* readers in their homes. Everyone interviewed had read the "current" issue of the *Post*, but had not, of course, seen the issue of the following day. The subjects were shown the advertisements in a folder, and asked:

"Will you please *identify the advertisements you remember having seen.*"

It will be noted that this form of the question puts the burden of proof on the individual. He is not asked whether he has seen an advertisement or not, but is asked only to identify or point out the advertisements he has seen.

At the close of the interview each person was asked whether he had *looked into* that week's *Collier's* or *Liberty*. Those advertisements which were to appear in the *Tuesday Post* and which had appeared in either *Collier's* or *Liberty* were then not included in the computation for those persons who had looked into either of these magazines.

Of the 277 advertisements remaining, 20 were double page, 60 were page colored, 94 were page black and white, 53 were half page, and 50 were quarter page. With one exception every advertisement which *had not yet appeared* was recognized as having been seen by some of the persons interviewed. The over-all average incorrect identification scores for the different types of advertisements are shown in this table by Lucas and Murphy: [3]

<i>Advertisements</i>	<i>Average Incorrect Identification</i>
Two Pages	20 Per Cent
Full Pages in Color	24 Per Cent
Full Pages in Black and White	20 Per Cent
One-half Pages	20 Per Cent
One-quarter Pages	17 Per Cent

More important than these averages, however, is the range of "false recognition," varying from 0 per cent for a full-page black and white Parke Davis to 58 per cent for a full-page colored Coca-Cola advertisement. Grouping the scores for all the advertisements we find that:

Range of Error in Recognition²

- 44 "ads" were falsely recognized by 10% or less of all people
- 96 "ads" were falsely recognized by 10% to 20% of all people
- 85 "ads" were falsely recognized by 20% to 30% of all people
- 30 "ads" were falsely recognized by 30% to 40% of all people
- 16 "ads" were falsely recognized by 40% to 50% of all people
- 9 "ads" were falsely recognized by 50% or more of all people.

This proof of the great variation in false recognition for different advertisements agrees with results from other experiments and indicates again the

² Adapted from Lucas & Murphy, note [3].

futility of attempting to "correct" for this factor, or even of assuming that it is a constant which may in some way or other be measured and allowed for. Therefore, if an advertiser by such a test gets a result showing that 70 per cent of the people have recognized his advertisement, he cannot tell whether 10 per cent or 50 per cent of the people were mistaken, *unless he has first made a test of this kind.*

A table showing this in dramatic fashion is reproduced herewith from Lucas: [2]

CORRELATIONS BETWEEN ORDINARY RECOGNITION SCORES AND CONTROLLED RECOGNITION SCORES FOR THE SAME MAGAZINE ADVERTISEMENTS*

Correlation Coefficients (Pearson r)

Recognition Scores

--by Leading Commercial Recognition Service

--and by Controlled Recognition Method

Saturday Evening Post Date (1939)	Number of Test Exhibits	Commercial Ratings Compared With			
		Uncorrected Recognition		Corrected (Controlled) Scores	
		r	P.E.	r	P.E.
March 18	23	+ .85	$\pm .04$	+ .66	$\pm .08$
25	20	+ .80	$\pm .05$	+ .56	$\pm .10$
April 1	20	+ .75	$\pm .07$	+ .61	$\pm .09$
Combined Three Issues	23	Full Pages in Color		+ .38	$\pm .12$

*There may be a spurious element in correlating a series of advertisements running from black half pages up to the more expensive colored pages. When the correlation is computed exclusively on the basis of colored pages in the three combined issues, it falls to +.38 as shown on the bottom line.

This table shows correlations between scores of commercial organizations and those obtained by Lucas with a scrapbook. The uncorrected scores for three consecutive issues of the *Saturday Evening Post* give correlations of +.85, +.80, and +.75. When the Lucas corrected scores are used the correlations drop to +.66, +.56, and +.61, respectively. This is further evidence that high or low uncorrected recognition scores do not predict comparably high or low corrected scores.

RECALL TESTS

Recall tests assume various forms, all of which utilize memory to an appreciable extent. However, recall tests are not tests of memory as such, but tests of the rememberability of particular advertisements or advertising campaigns. As opposed to recognition and identification tests, advertisements or test materials are not shown.

Unaided recall. Early tests of advertising (Gale, 1900; Scott, 1908) dealt mostly with unaided recall. A person who had read the current issue of the *Saturday Evening Post* was asked, for example, "What advertisements do you remember seeing?" This test gives results too general to be of use in solving most practical advertising problems.

Aided recall. A common variety of recall test is the aided recall test. An illustration of an aided recall question is: "What *brand of aspirin* have you seen or heard advertised recently?" The principal characteristic of this question is that it gives a measure of comparative brand dominance. It indicates the extent to which any particular brand is uppermost in the minds of the people interviewed. Obviously, factors other than advertising, such as the use or the prestige of a particular brand, influence the results. Because of this and other factors the relative dominance of the various brands does not necessarily correlate with their relative sales.

Here are the results, for example, of two studies made twelve months apart. [1] Each study was made in over 75 cities and towns, properly distributed by city size, geographic areas, 50 per cent men and 50 per cent women, properly distributed by socioeconomic groups. The question asked was: "What *brand of dentifrice* have you seen or heard advertised lately?"

<i>Brand</i>	<i>Oct. 1937</i> <i>Per Cent</i>	<i>Oct. 1938</i> <i>Per Cent</i>
Pepsodent	39.4	25.7
Ipana	18.4	18.6
Dr. Lyon's	13.7	17.4
Colgate	6.2	16.7
Listerine	4.2	4.9
Kolynos	3.5	4.1
Etc.	—	—
Total	100.0	100.0
Total Interviews	5,000	5,000

Two points are of significance here. One is the terrific loss for Pepsodent—a drop from 39.4 per cent in 1937, twice the number of votes received by the closest competitor, to 25.7 per cent in 1938. A special analysis showed practically all of this drop due to a loss of radio impact. The second change of significance is the gain made by Colgate—from 6.2 per cent in 1937 to 16.7 per cent in 1938. This gain for Colgate, interestingly enough, was made by radio.

The further question in reference to the brand named, "What did this advertising say?" is often asked. The answers give a picture of the specific advertising impressions which have been made, although here again experience with the product and other factors influence the answers somewhat. For example, here are some of the responses made by the men and women who answered "Pepsodent" above:

COPY TESTING

125

<i>Features of Pepsodent Advertising Recalled</i>	<i>Oct. 1937 Per Cent</i>	<i>Oct. 1938 Per Cent</i>
Brush your teeth twice a day (and/or) see your dentist twice a year	19.7	10.1
Contains Irium, new chemical, new ingredient	35.8	26.0
Gives you whiter teeth; keeps teeth white	10.9	11.6
Removes film; removes tartar	6.0	4.1
Keeps teeth clean; makes teeth brighter, makes them glisten; preserves teeth; good for teeth	6.9	16.0
Amos'n'Andy; radio program (no specific feature)	5.5	7.0
Etc.		
Total Per Cent	100.0	100.0

At a glance one can see what ideas Pepsodent has succeeded in leaving with people. Several points are of interest. First of all, Pepsodent had evidently succeeded well in "selling" the idea of Irium—of a *new* chemical or ingredient—by October, 1937. By 1938 this idea, although still outstanding, had slipped a little. Again, it may seem surprising that the "Brush your teeth twice a day; see your dentist twice a year" idea was so strongly entrenched in 1937—some time after Pepsodent had stopped using this theme in an aggressive way and on a large scale. This is testimony to the potency of the idea itself, evidence of which is available from many sources. The drop of the "Brush your teeth twice a day; etc." idea from 19.7 per cent in 1937 to 10.1 per cent in 1938 follows a well-established principle in psychology and advertising; namely, that no matter how good an idea or how strong an association, reinforcing at intervals is necessary in order to keep it alive.

A strikingly different pattern is observed for Dr. Lyon's powder. Here the responses were:

<i>Features of Dr. Lyon's Advertising Recalled</i>	<i>Oct. 1937 Per Cent</i>	<i>Oct. 1938 Per Cent</i>
Clean your teeth with powder—do as your dentist does	69.6	75.4
Gives whiter teeth; keeps teeth white	8.0	4.5
Keeps teeth clean; makes teeth brighter; makes teeth sparkle	5.6	4.2
Use powder (instead of paste); powder is best	7.3	5.5
Etc.		
Total Per Cent	100.0	100.0

Here we find mostly just one idea—"Clean your teeth with powder—do as your dentist does." Certainly Dr. Lyon's succeeded in impressing people with it. Triple Associates Tests made from 1932 to 1938 show a steady increase in the penetration of this idea. The preceding table on Pepsodent features remembered indicates the "tenacity" of the "dentist" idea. An examination of the sales records of Pepsodent during this time reveals a generally downward

trend. Dr. Lyon's, by contrast, showed a steady gain of customers during the same period.

Triple associates. This is a special form of recall test developed by Link [4] and is illustrated by the question: "What company advertises 'Better Things for Better Living Through Chemistry?'" This test can be applied only to those situations where an advertiser has used a central idea to carry his advertising message. Under these circumstances it gives an excellent picture of the extent to which the advertiser has succeeded in identifying this central copy theme with his brand.

Several years ago the Psychological Corporation was asked to test a newspaper campaign of a soap used primarily for dishwashing. The principal features of the campaign were: A company representative would make arrangements in a town for one of the local home economics experts to give a lecture on hygiene, cleanliness, etc., to a group of local women, some church or club group. At the end of the lecture, the company representative would put on a "scientific" demonstration of the effectiveness of the brand of soap in question, compared to other "leading" brands. This test consisted essentially in putting a tablespoonful of each of the soaps in a quart jar. One cup of boiling water was then poured into each jar. Invariably the advertiser's soap produced more and deeper suds than did any of the other "leading" brands. A local photographer then took pictures of this scientific demonstration, including some of the more prominent local women. This was then run in the local newspaper under the headline and story of "The Deeper Suds Test."

The first approach was to make a quick Controlled Opinion Test. Several variations of the test indicated that the advertisements interested most housewives only slightly. Of more importance was the finding that when the name of the advertiser was blocked out, there was considerable confusion with Super Suds, a competitive brand. On the basis of these findings it was recommended that the theme "The Deeper Suds Test" be changed, but the client was not convinced on the basis of the Controlled Opinion Tests and asked for further tests.

The following procedure was carried out. Four cities were selected for an eight-week newspaper campaign. Immediately preceding the advertising 300 housewives of the large middle socioeconomic group were interviewed in each of the four cities. Interviews were widely scattered throughout each city and only women in even-numbered homes and apartments were interviewed. Two questions were asked: (1) "What brand of soap do you use for dishwashing?" (2) A triple associates question, "What brand of dishwashing soap advertises *The Deeper Suds Test*?" The results, as can be seen from the table below, were nothing short of startling and definitely confirmatory of the Controlled Opinion Tests.

What brand of dishwashing soap advertises: "The Deeper Suds Test"?

	<i>Before Advertising Campaign Per Cent</i>	<i>After Advertising Campaign Per Cent</i>
Super Suds (Incorrect)	12	16
X—The Advertiser!	1	2
All other brands	3	3
Don't know	84	79
Total Per Cent	100	100
Total Interviews	300	300

The client refused, however, to be dissuaded. So after eight weeks of newspaper advertising the test was repeated, with women across the street and across the hall, that is, with women in odd-numbered homes and apartments. In this way the two samples were kept very comparable. The results indicated further that the company was doing an excellent advertising job for Super Suds! No further comment seems necessary.

This illustration indicates several things. First of all, it shows how the triple associates technique can be used prior to any advertising in order to discover the degree of confusion, if any, which exists between the central idea or theme proposed and other advertising which may already have strong claims on the idea. Secondly, it indicates the small amount of "guessing" that can be observed on the part of the large majority of the people who were uncertain and said they did not know. Most triple associates tests indicate a limited amount of confusion.

Reliability. Recall tests show reliability coefficients of approximately +.90, comparable to the recognition and identification tests.

Validity. Inasmuch as recall tests measure the strength of associations set up, not only by the advertising, but by use and familiarity over a period of time, their validity is self demonstrated. The results or responses to the questions asked are in themselves evidence of the degree of establishment of these associations.

Practical applications or uses. The purpose of much advertising is to establish an association between a need or want and a particular brand of product. For example, "I need tooth paste—I'll buy Ipana." That advertiser who can establish this type of association most strongly will, theoretically, have predetermined the consumer toward the purchasing of his particular product—*other factors being equal*. Other factors are, however, very often not equal—*one cent sales, special merchandising offers, availability, are illustrations of factors which come between the establishment of an association and the actual purchase and use of a particular brand or product. Therefore, there need not be expected a high correlation between sales and these tests.*

Recall tests are of most value when repeated at periodic intervals so that an advertiser can determine not only his comparative ranking in relation to his principal competitors, but also any increase or loss of his standing. Furthermore, periodic studies will indicate the extent to which particular copy themes or ideas, not only of his advertising, but also of his competitors', are making headway or losing ground.

THE CONTROLLED OPINION TEST

The rating of objects or items on the basis of the opinion or judgment of the subject is fundamental to much experimentation in psychology. The Controlled Opinion Test, popularly known as the Consumer Jury Test, is based on consumers' opinions under controlled conditions. It consists essentially in the showing of two or more advertisements to a person and asking questions such as: Which advertisement interests you most? Which headline appeals most to you? Which illustration do you like best? What is the story this advertisement tells? etc. This is not a test of the rememberability of advertisements as are the recognition and recall tests. It is a direct measure of people's interests, likes and dislikes, the story or idea the advertisement conveys, etc.

Reliability. Strong [5], using imitative as well as actual advertisements, secured reliability coefficients ranging from +.84 to +.95 with groups of subjects as small as 25 in number. Newhall [6] reported 71 coefficients of correlation of agreement among various subgroups of 12 to 41 subjects. Excluding seven correlations between men and women on advertisements which differed in their appeals to the two sexes, the correlations can be described thus: all were positive, 91 per cent were above +.50, 80 per cent above +.60, 66 per cent above +.70, 55 per cent above +.80, 23 per cent above +.90.

Validity. Starch [7] reported positive correlations between test rankings and sales, coupon, or inquiry rankings ranging from +.40 to +1.00, with an average of +.79 for twenty series of advertisements. Borden and Lovekin [8] reported a study made on seventeen series of magazine advertisements. Each series contained from two to five advertisements which showed reliable differences in sales or numbers of inquiry returns. Tests were made with men and women from economic, social and educational groups which included consumers for the products advertised. They found that when the article advertised was limited in appeal, inquiry or sales rankings and controlled opinion rankings agreed in only 36 per cent of the cases. However, when the advertised article had wide appeal, controlled opinion rankings and inquiry or sales rankings agreed 92 per cent of the time. These differences point to the necessity of testing a representative consumer group.

Practical application and uses. One of the principal advantages of this test is that it can be used to test advertisements *before* they have run—before large sums of money have been spent. Rough layouts may be used, which makes "pre-testing" relatively inexpensive. In many instances, there are incidental benefits over and above the selection of a specific advertisement or advertising campaign. For example, only two out of one hundred people mentioned "hexylresorcinol" when asked to tell "the story" of an advertisement in which this was prominently featured in the headline. The copy elaborated further, giving the "scientific" implications. This was hardly the overwhelming impression an aggressive sales or advertising manager might wish for.

Several years ago a well-known company had advertised a canned food in expensive space, featuring a contest with ten first prizes of paid two-week trips to Hollywood. The small number of entrants led the advertiser to wonder regarding the reasons. A simple controlled opinion test indicated that 42 per cent of the women who used the canned food with some regularity had babies or young children and would have found it very difficult or impossible to go to Hollywood! Another 23 per cent didn't want to go without their husbands! A supplementary test with unmarried girls 15 to 20 years of age showed 90 per cent interested in going to Hollywood. But only 17 per cent had ever prepared a can of the food. This advertiser might well have saved from \$60,000 to \$100,000 by a simple controlled opinion test of the basic appeal—the unwanted trip to Hollywood!

Warning. This test will give invalid results if the advertisements contain themes or claims which elicit "socially unacceptable" responses. Results may not be valid, for example, if some advertisements emphasize sex or moral values or if infants are used as illustrations, etc. Many advertising men believe the test has limited value because it is "artificial," i.e., the respondents become advertising critics instead of normal readers of advertising. Like all tests, the value of the results obtained depends upon sincere co-operation and interest in the test on the part of the respondents.

INQUIRY TESTS

Clipping a coupon or responding to an offer buried in the copy is an overt response. Many advertising men feel that therefore inquiry tests are particularly valid, despite studies which have shown that respondents often are not typical of the consumer population.

Under the category of inquiry tests are included all tests based on advertisements which invite a request of some kind from the consumer. The offers are so various that it is possible to suggest only a few of the more common:

1. The product itself, either sample or regular size.

130 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

2. A booklet describing the product, how it can be purchased, its uses, etc.
3. A premium such as a picture, pocketknife, or some novelty.
4. A special bargain, such as a three-year subscription for the price of two.

These offers may be free, free with some purchase or order, free except for a small postage charge, they may involve a sale at the regular price as in mail order advertising, etc.

The form in which the offers are presented also varies. Some of the more common are:

1. A coupon to be clipped or torn out forming a regular part of a newspaper or magazine advertisement.
2. A buried offer, that is, an offer found in the copy itself.
3. An offer inviting a written request, a telephone call, or asking a person to present a coupon at a dealer's, etc.

Variables which affect inquiry returns. Obviously, there are a great many factors which influence the number and the rate of inquiry returns. On the following there seems to be at least general agreement. Some are, of course, elementary.

1. *Size.* All advertisements tested should be the same size.
2. *Position.* Advertisements should occupy the same space since page position and position in the paper or magazine influence pulling power.
3. *Same offer.* The appearance as well as the offer itself must be the same. If a coupon, it should be in the same place in the advertisement and identical, if possible, in appearance and wording. If a buried offer, the wording and relative position in the copy, that is, in the last paragraph or paragraph before the last, should be the same.
4. *Coupon or buried offer.* A coupon will usually bring more replies than a buried offer. For instance, children and professional coupon clippers will often send in coupons. Only those who read most of the advertisement will discover a buried offer.
5. *The medium used.* The medium in which the tests are made should be similar to the media in which the final campaign is going to be run. On the basis of previous experience, the tester should know that the findings of the test medium can be applied to those in which final copy will appear.
6. *Rate of return.* There is quite often a dropping off of inquiries by the time the series of advertisements comes to a close; that is, the fifth in a series of weekly advertisements in a local newspaper may produce fewer replies because the market has been milked dry, so to speak. In order to take this into account, the first of a series of advertisements is often run again at the end and a trend line established. Then the advertisements are rated on the basis of their relation to the trend line. [9]
7. *The time between advertisements.* The time between each advertisement

of a series and the following one must be adequate to insure that the bulk of inquiries can be attributed to the particular advertisement in question.

8. *Competition from news.* A test in newspapers particularly may suffer from the competition of unusual news such as the Normandy invasion, our recapture of the Philippines, etc.

9. *Seasonal influences.* A series of test advertisements may run during several months or more. During this time seasonal variations in weather or seasonal changes in demands or habits may occur.

10. *Competitive changes.* During the six or seven weeks of the test period market conditions may vary. In fact, it has been suggested that competitors might even introduce special sales in test cities in order to "help" an advertising test of another brand or company.

11. *Weather conditions.* The test of appeals for the selling of umbrellas might be considerably affected by a hailstorm or rainstorm of some magnitude. In the same way, cold weather might definitely affect the results of a test of antifreeze for automobiles, etc.

Split-run technique. Many but not all of the preceding variables can be controlled by means of the split-run technique, that is, simultaneously running two or more advertisements in comparable parts of a magazine's or a newspaper's circulation. For example, Advertisement A appears in half the circulation, and Advertisement B in the other half. The advertisements appear in the same page position, and appeal to approximately comparable economic groups. They run during the same weather conditions and compete with the same news or editorial stories. This maximum practical comparability of external factors insures that differences in number of returns are due to differences in the advertisements themselves. There are now a fairly large number of magazines and newspapers which make available this type of operation to their clients.

Practical application and uses. The most common use to which the coupon or inquiry test is put is to determine the comparative pulling power of two or more advertisements, two or more appeals, two or more headlines, etc., in preparation for the running of a campaign in more expensive space. Usually, the tests are made in magazines or newspapers of small circulation. Recently, however, the tendency to utilize the split-run facilities offered by a number of magazines and newspapers has tended to increase the circulation of the testing media.

Of all tests this is probably the one the amateur most often goes wrong on. In general, he believes that all he need do is to put two ads in two different newspapers and observe which produces the most replies. Sometimes, the amateur knows that the larger the circulation, the more replies, usually. He may know that there are such things as seasonal variations, etc., but very often

he oversimplifies. In this connection a table which shows the seasonal indices determined by three different studies might be instructive.

TABLES FOR SEASONAL CORRECTION

Month	Starch [10]	Falk [11]	Rudolph [12]	
	1918-1930	1914-1928	1923-1927	1928-1932
January	120	184	124	99
February	130	157	125	144
March	121	107	116	118
April	99	77	96	104
May	86	67	78	89
June	83	65	79	90
July	83	76	93	91
August	92	92	111	91
September	108	95	105	96
October	116	92	101	94
November	99	89	91	101
December	80	99	72	85
Average	100	100	100	100

It is quite apparent that indices for seasonal variations differ. Certainly, if Falk had used the October index of Starch, based on inquiries of a wide range of products, in correcting for his replies to advertising of plan-cut homes and materials for homes, he would have erred considerably.

Reliability. Reliability coefficients vary considerably, ranging from $+ .25$ or $+ .30$ on the one hand to about $+ .90$ on the other. Coefficients are often low when returns are very scattered; for example, when only three or four hundred returns are obtained from an advertisement with a circulation of 500,000. Under these circumstances, reliability coefficients between split halves or thirds tend to be low. If the returns are higher, then the coefficients of reliability usually rise. In general, it may be said that the inquiry returns show high reliability, usually ranging between $+ .60$ and $+ .85$.

Validity. Even the split-run technique does not, of course, insure the validity of inquiry returns. The correlations of coupon or inquiry returns from newspaper and magazine advertisements with sales results or with any other validating criterion vary tremendously.

Returns from mail-order advertisements, which require part or all the price of the product advertised, approximate and in some cases actually are sales. Here face validity and validity approach identity and the correlation between coupon returns and sales must approximate $+ 1.00$. Coupon or inquiry responses are often followed up by mail or personal sales efforts. Here, too, validity can easily be determined. By trial and error methods over a period of years seasoned advertising men have in many instances determined the type of offer, media, and other factors necessary to give inquiry returns high validity for a particular product.

Most advertising situations do not, however, lend themselves to this type of validation. Hence the evaluation of the results depends upon the personal skill and judgment of the interpreter and tester. The varying abilities of those who do inquiry testing account, at least in part, for the varying validity of these tests. Here resides, without doubt, one of the major reasons for the divided opinion on the usefulness and adequacy of inquiry tests.

SALES TESTS

Sales tests try to measure the effects of advertising on sales. Since the purpose of most advertising has been to sell, it is easy to understand the high regard in which the sales test has been held. The sales test seems offhand an easy test, but actually it is extremely difficult because there are so many other factors (as in the inquiry test) which influence sales. Sales tests are also much more costly than any other form of copy test. Nevertheless, this method has been widely used.

Test cities. The most common form of sales test is one which purports to measure the comparative sales effects of two or more advertising campaigns. Usually a small number of cities is selected, one or two cities for each of the campaigns to be tested. An additional city or two is often selected as a "control"; here either no advertising is run or the regular advertising is continued. Cities are usually selected for comparability in size, in average income of families, in newspaper circulation, in nationalities of the residents, in types of employment available and on the basis of other similar factors which affect sales. Other variables or influences which must either be controlled or taken into consideration are weather conditions, epidemics, seasonal influences, price and price changes, duration of test period, competitive factors, and special promotional efforts.³ Unless these critical variables are controlled, or their importance known, the sales test may well be of little value. Sales are measured before, during, and after the campaign, and the campaigns rated according to their comparative sales.

An example of a sales test. Giellerup wished to discover the best appeal and form of advertisement to use in selling a roofing paint. [14] He particularly wanted to compare:

1. "Regular" with "editorial" type advertisements.
2. Advertisements with logotype to those without.

Four cities, ranging from 59,000 to 85,500 in population were selected as test cities. The buying power per capita, the number of homes, the number of rented homes, the circulation of the newspapers used and other relevant facts were obtained for each of the four cities.

³ For a more detailed treatment see note [13].

Four advertisements were prepared and each run four times on the same days in each of the cities during an eleven-day period. An exact record of weather conditions was kept.

One of the advertisements was of the editorial type without logotype. The other three were the regular type of advertisement, one with logotype, and the other two having the name of the company at the bottom. An accurate store record of the sales was kept—extending for a period of two days after the last of the four insertions.

The results indicated that the editorial type of advertisement was definitely superior.

A second test. Using the results of the first test as a basis, four advertisements were made up, all with the headline "Put An End to Your Roof Troubles." Two of these advertisements were the editorial type, identical except that one had a logotype of the name of the paint at the bottom. The other two were more the regular type of advertisement, and identical except that one also had a logotype of the name of the paint at the bottom.

These four advertisements were run three times during a one-week period in the same towns as before—starting two and one half weeks after the finish of the first test.

Town	Popula- tion	Buying Power Per Capita	Circula- tion of Paper	Ad*	Test 1	Test 2	
					Sales Per Cent	Ad*	Sales Per Cent
A	60,667	576	7,574	R	11	R	14
B	85,545	2,490	16,304	R	8	Ed	46
C	59,184	1,048	6,377	Ed	62	Ed,L	27
D	77,270	894	12,776	R,L	19	R,L	13
				Total	100		100

*R, Regular type of advertisement.

R,L, Regular type of advertisement with logotype.

Ed, Editorial type of advertisement.

Ed,L, Editorial type of advertisement with logotype.

As a result of this sales test, Giellerup was definitely able to say for this campaign:

1. That "editorial" copy was better than "regular" copy.
2. That advertisements without logotype were superior to those with, aside from costing less.

Methods of measuring sales. Sales before, during, and after a sales test are measured in one or more of the following ways:

1. *Company sales.* Sales of the company to retailers, wholesalers, and jobbers are used as a basis for measuring the effects of sales tests. Obviously this is a crude means of measurement because of the unknown lag between sales by the company and sales to the consumer due to warehouse inventories, dealer inventories, uncertainty of areas serviced by specific jobbers, etc. At the present time this measuring device is seldom used.

2. *Sales of retail outlets.* A sample of retail outlets in the test and control cities (or on a national basis) is inventoried periodically to measure the flow of merchandise. An important contribution has been made here by the A. C. Nielsen Company in the drug, grocery, and liquor fields. Sales are measured by means of bimonthly inventories in a large number of retail outlets scattered throughout the United States. (The most important defect for some products is the refusal of certain chain outlets to co-operate.) Results are reported by brand in terms of units sold, and also in terms of dollar volume. Many companies use the Nielsen service for measuring results of their sales tests rather than setting up a separate sales checking system.

3. *Measurement of consumer purchases.* A recent trend has been to check sales tests by means of consumer panels or by means of repetitive studies in comparable groups of homes. The two most widely recognized regular commercial services are the National Consumer Panel by Industrial Surveys Company and the Psychological Corporation Brand Barometer studies—both described more in detail in special chapters in this book. In consumer panel studies arrangements are made to have regular records kept by family members of all purchases, or of all purchases of certain kinds (e.g., drugs, cosmetics, house furnishings, etc.). These records are then tabulated, analyzed, and the results interpreted accordingly. In the periodic Brand Barometer studies, questions on recent purchase or usage are asked in 10,000 urban homes scattered throughout the country and sales trends recorded accordingly. In both these procedures information of the advertising influences to which the family members have been exposed can be obtained, and more accurate analyses of advertising influences therefore made. Probably the most important characteristic of these modes of measurement is that they occur at the terminal point, the consumer, who is the person to whom most advertising is addressed. By these methods the effects of advertising can be measured in terms of socio-economic status, age, sex, educational level, etc.—all important to the advertiser but not obtainable by other modes of measurement.

SUMMARY

There are many variations of the six types of tests discussed above. In addition, there are other less commonly used methods of testing advertisements. It may safely be said of all forms of copy testing that no one is the final answer to the problems of the advertising man.

Most of these tests show fairly high reliability coefficients but their validity is much harder to demonstrate. However, test results often "make sense," not only to practical advertising men but also to scientifically trained research men.

Copy testing, like medicine and engineering, belongs to "the arts of prac-

tice." It has not attained scientific status. It may never achieve such status. Nevertheless, even in its present undeveloped and ill-defined state, copy testing gives valuable clues to advertising men, partly because many of the questions arising are of even a grosser nature than the copy testing tools employed to solve them.

REFERENCES

1. Unpublished studies available through the courtesy of the Psychological Corporation.
2. Lucas, D. B. A rigid technique for measuring the impression values of specific magazine advertisements. *Journal of Applied Psychology*, 1940, 24, 778-790.
3. Lucas, D. B., and Murphy, M. J. False identification of advertisements in recognition tests. *Journal of Applied Psychology*, 1939, 23, 264-269.
4. Link, H. C. A new method of testing advertising effectiveness. *Harvard Business Review*, 1933, 11, 165-177.
5. Strong, E. K. The relative merit of advertisements. *Archives of Psychology*, 1911, 17, 81.
6. Newhall, S. M. The reliability of the order of merit evaluation of advertisements. *Journal of Applied Psychology*, 1930, 14, 532-548.
7. Starch, Daniel. *Principles of Advertising*, 1923; Testing the effectiveness of advertisements. *Harvard Business Review*, 1923, 2, 464-475.
8. Borden, N. H., and Lovékin, O. S. A testing of the consumer jury method of ranking advertisements. *Publication of the Graduate School of Business Administration*, Harvard University, 1935, 22, No. 2, *Business Research Studies*, No. 11.
9. Caples, John. *Tested Advertising Methods*, 1932; *Advertising for Immediate Sales*, 1936.
10. Starch, Daniel. *An Analysis of Five Million Inquiries*, 1930.
11. Rudolph, H. J. *Four Million Inquiries from Magazine Advertising*, 1936.
12. Falk, A. T. Analyzing advertising results. *Harvard Business Review*, 1929, 7, 185-194, 312-329.
13. Coutant, Frank. How to select trial cities and check sales-test results. *Printers' Ink*, 1932, 161, 59-62.
14. Giellerup, S. H. Unpublished study made available through the courtesy of S. H. Giellerup, Partner, Marschalk and Pratt Company.

* * * * *

Firth, L. E. *Testing Advertisements*. New York: McGraw-Hill, 1934. Discussion of Copy Tests, p. 282.

The Advertising Research Foundation. *Copy Testing*. New York: Ronald Press, 1939. Discussion of Various Tests, p. 131.

Chapter X

THE OFFICE OF RADIO RESEARCH A Division of the Bureau of Applied Social Research Columbia University

MARJORIE FISKE AND PAUL F. LAZARFELD
Office of Radio Research

THE rapidly increasing development of the radio industry in the past two decades has opened up a new and increasingly important area of communication. Radio is accessible to more people than any other kind of communication. Its effects are therefore of increasing importance to the sociologist, the educator, and the politician. And since radio in America is a privately owned and operated industry, its impact is also a matter of importance to networks, advertisers, advertising agencies, and others concerned with its commercial effectiveness.

The Office of Radio Research has been functioning for seven years, during which it has been concerned with a wide variety of radio research problems. In some instances these problems could be solved by a relatively simple adaptation of techniques used in other fields of research. In others it was necessary to develop new techniques to fit the special characteristics of this new medium. It will not be possible within the scope of this chapter to enumerate all such adaptations and innovations. We shall, rather, touch upon a few which illustrate the interrelationship of radio and other fields of communications research.¹

THE OBJECTIVES OF RADIO RESEARCH

It is manifestly impossible to study either the content or the effect of every radio broadcast that goes on the air. The larger body of knowledge, the overall picture, has to be built up over a period of time, segment upon segment,

¹ The authors are indebted to Dr. Bernard Meyers for his assistance in organizing this material.

each segment representing a study of a particular program or a particular group of listeners. Furthermore, studies of particular programs or particular groups of listeners may be done from the different standpoints of the educator, the politician, the sociologist, the psychologist, and the businessman, each seeking answers to different questions. Rarely is one program or one group of listeners studied from all of these viewpoints, but each study contributes to the sum total of knowledge of radio's role in our culture. Each contributes new techniques which can be used by the other. Altogether, they are gradually weaving together to form an increasingly important part in the general pattern of communications research.

Some examples of the several approaches to radio research. The Office of Radio Research has had occasion to do research from all of these several standpoints. In studies reported in *Radio and the Printed Page* [1] radio was viewed from the standpoint of the educator; reading and listening habits were compared and new insights gained as to the relative roles of radio and print as informational media. It has studied a particular program, the Orson Welles "Invasion from Mars" broadcast, for example, from the standpoint of the psychologist to gain insight into how different kinds of people react to a given stimulus situation. [2] This broadcast, it will be recalled, resulted in near panic in certain parts of the country. To determine how it came about that a radio drama could spread genuine terror through a substantial part of the population, detailed interviews were made with many different kinds of listeners who reacted in a variety of ways. The results not only shed light on the power and potentialities of radio as a medium of communication, but enabled the psychologist to gain insight into the general psychology of panic.

Or, again, certain programs and certain kinds of listeners have been studied from a broad sociological standpoint, as in the not yet published study of the Kate Smith all-day bond-selling appeal, wherein a particular broadcast of a popular entertainer was subjected to scrutiny. In this study the content of the program was analyzed to determine the variety of its appeals, and the listeners were interviewed at length to determine the relative impact of these appeals, and how they came to decide to buy a bond from Kate that day. In the course of the analysis it became apparent that a particular radio entertainer may epitomize the social force of radio, reflecting certain trends and concepts in our culture, and at the same time reinforcing them. Kate, for example, lays great stress on the sacrifices and sacredness of motherhood. Indeed, for many of her listeners she has come to typify motherhood: "Only a mother could plead the way she does," even though most of them know she is not a mother. In this extolling of motherhood she not only reflects one of the basic concepts of our time and our culture, but at the same time she reinforces and strengthens it.

The program sponsor and the advertising agency study radio from yet another angle. They are concerned with the extent to which their "messages" get across, and with the extent of acceptance or rejection of their particular programs. Here the research focuses on the immediate response reactions of the listeners. An effort is made to determine the extent to which a program is liked or disliked and why, and its effect on the subsequent attitudes or actions of those who heard it. Studies of this kind involve either program research or the testing of commercials, techniques which we shall consider shortly.

Viewing radio as a whole, it is apparent that its content is the product of contemporary culture and customs. Its content reflects this culture because the people responsible for it are in turn the products of it. *Analyses of the content* of radio broadcasts, therefore, provide the sociologist with a better understanding of society. On the other hand, this content has an impact on millions of radio listeners, and by studying the effects of radio on listener's habits and attitudes the sociologist also gains insight into the way in which it is changing, modifying, or reinforcing the cultural pattern.

THE TECHNIQUES OF RADIO RESEARCH

Surveys of general listening habits. The most general kind of listener survey involves a simple count of how many people listen to a given program. Such counts, known as program ratings, are based on careful samplings of the population, and are made systematically by a number of research organizations. Program ratings and fluctuations are thus made available regularly to private clients.² The Office of Radio Research, however, while it does make use of such ratings in some of its more quantitative studies, e.g., *The Social Stratification of the Radio Audience*, by Hugh Beville, [3] confines itself largely to studies of a more detailed psychological nature.

Surveys of general listening habits may be made for several reasons. One may wish to compare the influence of the various media of communication on a given area of behavior. How, for example, does the influence of radio compare with the influence of newspapers and magazines on voting behavior? [4] Or, one may want to measure changes in listening habits resulting from program changes or to gain insight into the role of radio among certain groups of people. In this case one must also survey general listening habits over a period of time. Whatever the purpose of such general surveys, the procedure is the same: something akin to a "listening diary" must be procured from a representative sample of the population one wants to study.

Several studies of this kind, both commercial and noncommercial, have

² See Chaps. XI, XII, XIII, and XIV.

been made at the Office of Radio Research. A good example of this is the question of listening in the daytime. If the whole listening pattern of women were taken into consideration, they would fall into three large groups. Daytime listeners include those who listen to serials, those who listen to other programs and others who because they are at home in the morning could listen but do not. Thus radio actually does not reach a third of the available morning audience and, at the same time, has to cater to two rather different kinds of audiences. How to reconcile the interests of these two divergent sectors of the audience is a problem that leads to a large number of interesting and still partly unsolved research problems.

Another survey of general listening habits was concerned with the question of who listens to small local stations. This involved interviewing a cross section of the radio audience in a given locality about their radio listenership over a period of time, and comparing different age, sex, and socioeconomic groups in respect to station preferences. [5] It developed that people on the lower socioeconomic strata tend to listen to such stations more than do those who are better educated and better situated financially.

A study of a somewhat different nature recently completed by the Office of Radio Research [6] also falls into this category. Here the problem was to determine (a) the degree of satisfaction with current radio offerings, (b) attitudes toward commercial advertising on the radio, and (c) general receptivity toward a proposed new plan by which the listener would subscribe to a service which would provide him with three types of radio programs without any advertising.

Still another kind of general listenership survey is designed to determine the role of radio in the lives of particular groups—children, for example, or housewives, or certain socioeconomic groups. Such a study usually involves careful and detailed case histories of representatives of the group under study. This kind of investigation is well exemplified in two Office of Radio Research studies, *Listeners Appraise a College Station* [7], and "Radio Comes to the Farmer" [8]. In the latter, it was possible, by use of the detailed interview method, to determine the extent to which the acquisition of a radio changed the habit and thought patterns of a group of Iowa farm households.

A similar type of survey is indicated when a sponsor or a group of sponsors wants to measure the impact of his radio appeals, or to compare it with appeals in other media. It was found possible in one study, for example, to gauge both the actual results of radio advertising and to get some idea of its potentialities. [9] The investigators went into a representative, moderate-sized community and interviewed several hundred housewives at great length about their listening habits, their awareness of retail merchandising over the air, their varying degrees of receptiveness toward the various kinds of retail

advertising and the extent to which such attitudes influence their buying habits. Among other things, this study indicated that certain kinds of programs are better adapted to the selling of retail merchandise than others and that certain kinds of merchandise lend themselves better than others to advertising over the air.

In the course of such studies it has become clear that nonlisteners are also an important factor in radio research, both from the standpoint of particular programs [16] and from the standpoint of nonlistening in general. If we know why certain people do listen to a given type of program and how many people do not listen for these same reasons, we can plan program changes that may not only improve content level but at the same time increase the total amount of listenership. Similarly, if an extensive survey were made of people who seldom or rarely listen to the radio at all, we could round out our picture of radio as a cultural expression and a cultural tool.

The nature of program research. Such broad audience surveys as those outlined above cannot possibly encompass the more specific problems of listener likes and dislikes, listener gratifications or the extent to which listener attitudes are changed or modified by radio listening. Therefore, the investigator finds it increasingly necessary to study particular programs or series of particular programs. In doing so, however, he comes face to face with research problems which are not especially peculiar to radio but have their counterparts elsewhere in the communications field. How do you measure listener reaction to a program? What specifically does a listener mean when he says he liked or disliked a certain program? How can one measure the "effectiveness" of a given informational program? How determine the cumulative effect of a series of programs?

There are three different ways of learning what a program means to people: by subjecting the program to a content analysis, by making a differential analysis of the personal characteristics of the groups that listen to the program, or by asking people directly what the program means to them. Whenever possible, all three methods should be used simultaneously.

Content analysis of radio material involves essentially the same techniques as are used in the analysis of printed materials and are usually based on scripts or transcripts of the broadcasts (unless one is studying all programs for the occurrence of a certain type of content in which case one has to resort to "monitoring"). From analysis the investigator is able to list most of the affective factors of a broadcast. Thus the content analyst, after listening to a few instalments of a daytime serial script, may learn that it stresses an individualistic, competitive type of social relationship, that the surgeon hero wants to be a great man and stand high in a prestige hierarchy, that his interest is in himself and not in humanity. Or he may discover that the Negro

character depicted in the series is a servant whose chief characteristic is dog-like devotion to his master with little or no portrayal of any individual thoughts, feelings, or individuality of his own. In another study a content analysis of a Kate Smith script reveals that she sometimes uses the word "America" or "American" as many as seven times in five minutes, thus building up an associative complex that contributes to her reputation as a patriot.

But, as we have already suggested, content analysis is important largely as the first step in the study of any particular program or series of programs. In subjecting the script or scripts to a preliminary content analysis, the investigator accomplishes two objectives. He is able to distribute his questions on the various components of the broadcast with some regard to their frequency and importance. Secondly, a thoroughgoing content analysis permits certain inferences about what the listeners may get out of the content, or at least will give the investigator some idea of what *not* to look for. Because it provides both balance and perspective, content analysis is usually the first step in program research, whether it be an investigation of one program or of a series of programs.

The second way to find out what a program means to people is to determine what sex, age, and social groups listen to it. Much is known about the psychological differences among various strata of the population, and if the program is listened to by some of this group more than by others, the nature of its appeal can be more readily understood. If, for example, the audience of one of two comedians is more highly educated than the audience of another, then it can safely be assumed that the first offers a more sophisticated kind of humor. The characteristics to be isolated will of course vary with the problem at hand. In a study of the audience for a child guidance program, for example, whether or not the listener has children was a *pertinent factor*. Quite a number of childless women were found among the regular listeners, hence the conclusion that the practical advice offered is not the only appeal in this program. Some women, regretting their lack of children, derived a vicarious satisfaction from hearing child problems discussed, while for others the broadcasts had general educational value.

If more general listening habit surveys included detailed information about the listener, such as reading habits, leisure-time activities, community participation, and so on, this material would become a useful tool for the further analysis of what certain kinds of programs mean to listeners.

One of the major problems of program research is how to get the respondent to indicate what, in the program or series of programs under study, is responsible for his reactions to it. It means little or nothing to the program planner, for instance, if he is told that 70 per cent of the respondents studied

liked a program very much, 20 per cent liked it moderately well, and 10 per cent did not like it at all. It does not tell him what could be done to improve the attitudes of the other 30 per cent or whether changing it to meet their taste would not at the same time antagonize the 70 per cent who liked it in its original form. After considerable experimentation, however, the Office of Radio Research has developed a technique that seems to contribute to the solution of this basic problem. This technique involves the adaptation of the polygraph frequently used in experimental psychology, and is known as the Lazarsfeld-Stanton Program Analyzer.

The Program Analyzer is an apparatus that enables a group of respondents to record their reactions to a radio program as they listen to it, by pressing red ("dislike") and green ("like") buttons, or by not pressing buttons, which signifies indifference to what is being heard. The push buttons are connected with pens that move along a roll of tape synchronized with the radio program, thus making a permanent record of the reactions of all members of the group. [10]

Such a record alone, while interesting as a picture of the high and low points of the program as far as a given group of listeners is concerned, is comparatively meaningless from the standpoint of program improvement. The important thing for the program planner or the educator to know is what it was about a particular part of the program that the listener found dull or interesting and *why*—in terms of the listener's own background and experience. The Program Analyzer Technique is therefore nearly always combined with a *focused interview*³ in which the trained investigator, using the Program Analyzer graph for reference, is able to determine just what it was about the program that caused the reactions indicated on it, and just what these reactions mean in the experience of the listener. Every research man who has tried to determine the "why" of reactions to a particular experience will recognize the advantage of this method. It gives him a picture of reactions that occurred simultaneously with the experience, and obviates a frequent difficulty in retrospective interviewing: to wit, that the respondents often fail to remember how they felt in the earlier parts of the experience. Like many techniques developed in one field of communications,

³ The "focused interview" is a term applied to the technique of determining reactions to a particular communication or experience (a motion picture, a radio program, printed material and so on), which is known to the investigator—as distinguished from the more diffuse type of interview which is required when studying listening habits or attitudes which may be the result of several different experiences which are usually unknown to the investigator. The focused interview is a rather complicated procedure, and the O.R.R. is now in the process of codifying the results of its experience with this technique with various media of communication. The results of this systematization will be available in the form of a handbook sometime in 1946.

this one is useful in others as well and has been used successfully by the Bureau of Applied Social Research in testing reactions to motion pictures.

The Program Analyzer, of course, can be used to study reactions to any kind of program. It has been found useful in determining the effectiveness of educational broadcasts, in analyzing the appeal of entertainment programs, and in measuring the impact of commercial announcements. The usual procedure is to interview 10 or 20 groups of people (10 to 15 in a group) carefully selected to be representative of the audience the program is designed to reach. They first listen to the program, recording their reactions with the Program Analyzer push buttons, and are then questioned by a highly trained interviewer with all remarks recorded by a stenotypist. Their comments are then analyzed in conjunction with Program Analyzer graphs, and the investigator can thus determine what the effective components of the broadcast were and can make recommendations as to which parts of the program should be changed, or eliminated for more satisfactory results.

When a radio investigator wants to probe deeper, to determine the gratifications of certain segments of the radio audience, interviews of a more elaborate kind are in order. Such studies usually involve two steps: detailed and exploratory case studies, followed by less detailed interviews with a larger sample, for statistical verification of the hypotheses developed from the qualitative data. This combination of qualitative and quantitative research has two advantages. On the one hand, the first step enables the investigator to gain rich psychological insights which enable him to cover the wide range of possible responses in the statistical survey. On the other hand, the qualitative material enables him to understand, clarify, and illustrate the quantitative data more fully. The combination of the two types of research has proved so fruitful that it has become established procedure in many of the studies undertaken by the Office of Radio Research. Perhaps the best way to illustrate its value is by way of a concrete example.

An example of program research. The problem was to determine the gratifications of the millions of women who listen to the serial stories broadcast throughout the day by the major networks. As a first step, 100 women from various age and socioeconomic groups were interviewed intensively. [11] Analysis of their reports about their listening experiences and the satisfactions they derive from it indicated that there are three major types of gratification in listening to daytime serials. Some listeners enjoyed them primarily as a kind of emotional release. Burdened with their own problems, they claimed it "made them feel better to know that other people have troubles, too." A second and more obvious form of enjoyment of the serials comes from the vicarious experiences they supply. A third gratification was entirely unanticipated by the investigator and constitutes a good illustration

of the value of this kind of intensive interviewing. *Many* women listened to serials because they provide standards of value and judgment, and help them to solve everyday problems. They learn things from these stories which they use later in solving their own problems: "Bess Johnson shows you how to handle children. She handles all ages. Most mothers slap their children. She deprives them of something. That is better. I use what she does with my own children." Or they provide comfortable philosophy for use with one's self or others: "When Clifford's wife died in childbirth the advice Paul gave him I used for my nephew when his wife died."

In this way, by providing such "leads," the intensive interviews opened up the areas for investigation on a more quantitative basis. Later, when 2,500 listeners were interviewed [12], 41 per cent claimed to have been helped by daytime serials, thus giving statistical validity to a gratification that might have been overlooked altogether had the intensive interviews not been made. With the larger sample it became possible to make cross-tabulations which showed what kind of women found the serials helpful in this way. Thus, for example, it developed that the less formal education a woman has the more help she derives from the serials. The quantitative material also made it possible to analyze the nature of this help, and it developed that listeners find these programs useful in several ways: getting along with people, helping people with their personal problems, learning how to handle themselves in particular situations, learning how to accept misfortune with a smile, and so on.

Analysis of listener reaction combined with content analysis [13] of the scripts themselves then led to certain inferences about the role of such programs in our culture. It was found, for example, that these so-called true life stories do not deal with basic social or economic problems. They do not show a woman how she can improve her economic status, nor do they give her a better understanding of the current problems of our time—such as those of minority groups. They tend, rather, to imbue the listeners with a fatalistic philosophy of life: this is how it is, we aren't as badly off as we might be. They help the listener to accept her fate by universalizing it—"husbands never understand their wives." Thirdly, they encourage the listener to live life through ready-made formulas for behavior rather than helping her to develop a critical sense that will enable her to determine what is good or bad for her in a particular situation.

The field of program research, however, has been by no means completely explored. Little, for example, is known about the maximum potential of radio from an educational and cultural standpoint. We know, to be sure, that, by and large, the programs that are known and promoted as "educational" reach a relatively small proportion of the radio audience, chiefly those who

would make a point of acquiring the same information from another medium if it were not available to them over the air. [14] It is known that such programs will not reach even these relatively few listeners unless organized efforts are made to "build" an audience. [15] But what about the utilization of such already accepted programs as the daytime serials as a means of raising, rather than catering to, the cultural level of the average listener? The sponsor feels he would thereby lose some of his audience. But the fact remains that few have tried to improve them and there is as yet no proof that the sponsors are right or wrong.

Effect Studies. Another, more specialized form of radio research pertains to the effectiveness of one section or element of a program. The commercial sponsor may want to determine the effectiveness of his commercial announcement. He may want to compare the effectiveness of two or more different presentations. The program planner may want to determine the extent to which his program depends upon the popularity of any single feature in it. He may want to compare commentators or announcers to determine which one is most acceptable to the greatest number of listeners. In all these cases the research procedure, as in most stimulus-response studies, involves the holding of all factors constant except the one under study. Thus, to determine the relative appeal of two commercials, matched groups of respondents (or sometimes the same respondents) will listen to two broadcasts which are alike in all respects except the commercial. If there are no extraneous factors involved, all differences in reaction to the two will be the result of differences in the appeal of the two commercials.

If the sponsor does not have two or more specific commercials which he wants to compare, but wants, rather, to determine the effectiveness of a particular one, the problem is somewhat different. In the first place, he must decide whether he wants to measure the effectiveness of the commercial in terms of the number of sales of his product which it induces or is likely to induce, or whether he is concerned only with the extent to which the commercial is liked or disliked. (The relationship between liking a commercial or any other kind of persuasive appeal and being induced to act as a result of it is, incidentally, a problem which needs much further exploration. Studies done to date indicate that one may dislike a commercial intensely—"spot" announcements, for example, or singing commercials—and still be influenced by them.) If the investigator is primarily interested in sales effects rather than in what elements of the commercial make it effective, a controlled check is commonly used. A section of the population is "exposed" to the appeal and sales figures for the product in that area are checked against those in a comparable area where the population was not so exposed. An alternative to this procedure involves interviewing buyers of the product to

determine how they came to buy it. Most advertisers, however, seem to operate on the theory that there is a connection between liking a commercial and buying the product it extols. Consequently, they are interested in research that will determine the degree of acceptance or rejection of the commercial announcement itself.

If the advertiser is concerned only with the interest aroused by the commercial in a given program context, the *Program Analyzer technique* is in order. The graph will show clearly the relative position of the commercial within the framework of reactions to the program as a whole. From the focused interview that follows he can then learn much about what was liked or disliked about the commercial and what in terms of the listeners' own experiences caused the favorable or unfavorable reaction. This technique is useful also in determining the effectiveness of commercials placed at various stages of the program—e.g., is the commercial placed at the beginning, end or middle of the program more effective? Should it follow a peak of interest in the program to capitalize on the high degree of attention at that point, or would such an approach cause a "letdown" on the part of the audience which might boomerang with resentment that "something is being put over" on the audience?

Another technique for studying commercial announcements has been found especially useful in testing reactions to "touchy" subjects, very personal products, or in testing institutional advertising. This involves an intensive "depth" interview, which is made immediately after the subject has read or heard the advertisement, and is equally useful for printed advertisements. Here the interview is customarily of an associative nature. What words or ideas are taboo? What words cause unpleasant associations which might in turn result in an unfavorable attitude toward the product or sponsor? This technique, incidentally, has suggested the interesting possibility that certain matters can be discussed in print which are not acceptable over the air and that, contrariwise, some approaches are more effective orally than in print.

By and large, however, it has been found that the best way to make people articulate about commercial announcements, a subject that often leaves them lethargic at best, is to have them compare two or more. Most people are not sufficiently interested in such matters to become very talkative about their reactions, and the necessity of making a choice between two or more often provides the necessary impetus to self-examination as to why they selected one or the other.

Another type of program research problem to which we have already alluded involves the program series. To investigate just one of a series of educational or entertainment or dramatic programs would not be a valid test of the effectiveness of the series, for reactions to a given program may be

partly predicated on remembrance of what went before and expectation of what is to come. Then, too, if the investigator is interested in changes of attitudes as a result of a program series he will get little from merely testing one program. The technique that has been developed to solve this problem is called the "panel," [14] which, reduced to its barest essentials, involves the selection of a group of people who agree to listen to a series of broadcasts and then report their reactions to the various programs. They may agree to come to a given place in a group and participate in a group interview using the Program Analyzer, or they may agree to record their reactions to various programs on formal questionnaires (the latter is, of course, in order when a nation-wide sampling is desired). Thus a virtually constant and identical group is made available for the examination of a series of programs, and detailed comparisons of one program with another in a series can be made. Obviously, one can also procure other information from such a group—reading habits, program preferences, movie attendance, and so on, which is helpful in evaluating variations in listener reactions.

Special characteristics of radio. There are at least six distinguishing characteristics of radio which have to be taken into consideration when studying its effectiveness from any particular standpoint. Each of these qualities has both positive and negative aspects from the standpoint of effective communication.

Perhaps most significant of these characteristics is radio's *accessibility*. Nearly every person in the United States has access to a radio. There are few geographical or economic barriers to its use, once the initial investment has been made. In a sense, then, radio is more readily available than the other mass media, for each magazine, newspaper, and motion picture must be purchased separately. But in another sense radio is less accessible than these other media. Once one has bought a newspaper or magazine one can keep it. It may be read at any time and an interruption or a lack of comprehension of a passage is not a serious matter, for it can always be reread. But a radio program is as ephemeral as time itself. If the telephone or doorbell rings just when John Kieran is about to answer a baseball question or just when the mystery is about to be solved, one cannot set back the needle to pick up what has been lost. Motion pictures also are ephemeral in this sense, but the circumstances under which they are seen tend to offset this factor: one is not likely to be interrupted in a theater, and if a moviegoer so desires, he can always sit through a second showing.

Another special characteristic of radio is that it relies on *auditory perception*. The human voice is a more personal, direct, and potentially more stimulating means of communication than the printed word, but this does not necessarily mean that radio is a more effective means of transmitting all kinds

of communication to all kinds of persons. Studies done by the O.R.R., for example, indicate that people in the higher social and economic brackets more often prefer to read factual information rather than to hear it.

A third characteristic of radio is in part the outgrowth of the first two. Its accessibility combined with its reliance on auditory perception enables people to listen while carrying on a variety of other activities which do not necessarily interfere with their perception. But at the same time this quality of *noninterference* leaves the radio program liable to a low degree of attention. The listener may become so conditioned to it that he no longer hears it with any degree of acuity. This poses a problem for the investigator and necessitates a thorough probing of the seemingly factual statement: "Yes, I listened to that program" to determine the degree of concentration concealed behind it.⁴

A fourth characteristic of radio is that it *continues in time*. This means that a series of programs may become part of the daily or weekly habit patterns of the listeners, that *cumulative effects* can be built up over long or short periods. But it also means that it is liable to surfeit. It may be true that "if you hear a thing often enough you will come to believe it," but it is probably equally true that if you hear a thing too much you may not pay any attention to it at all after a time. Just where repetition ceases to be effective, just when saturation points are reached, is still a problem that has to be faced anew for each kind of program or message.

There are two other characteristics of radio which develop from its accessibility. A national network may reach into homes all over the country, but if it does so it must *confine its appeal to a general one*. This national quality prevents it from appealing directly to local interests and experiences. Theoretically, of course, the potential audience is great enough for a nation-wide program to be beamed at special groups such as *fishermen, students, or stamp collectors* and still reach a sizable number of people. But since the aim of the networks is to reach as many people as possible at a given time, their specialized appeals are confined to large groups such as farmers or housewives who are known to constitute the majority of listeners at certain times of the day. Appeals to smaller groups are left for the local stations which cannot hope to compete with the networks on their own ground. The coming of frequency modulation might bring many changes in this respect.

The discussion of the nature of the problems met in the field of research and some of the techniques developed to meet them may be sufficient to bring the reader to two conclusions already evident to social scientists con-

⁴The problem of developing techniques to gauge degree of attention to radio programs is becoming a matter of great concern to television producers, who want to know not only whether and how well a program was heard, but also whether and with what degree of attention it was seen.

cerned with radio. First, there is a growing awareness of the necessity of systematizing the knowledge and experience accumulating in this field: a conviction that such self-conscious rigorization of procedures will be of value not only in the field of radio research alone, but to the science of communications in general (and perhaps even to other fields of social research). Secondly, as this formulation and formalization of procedures and problems proceeds, the sociologist and psychologist working in radio research become increasingly humble about what they do not know. But even at this comparatively early stage of its development, radio research activities have already stimulated magazine and newspaper publishers to do more research than ever before, and it is possible that in the not too distant future not only will techniques and problems be exchanged between these two fields, but funds and research institutions may be merged for the greater benefit of both.

REFERENCES

1. Lazarsfeld, Paul F. *Radio and the Printed Page*. Duell, Sloan and Pearce, 1940.
2. Cantril, Hadley, Herzog, Herta, and Gaudet, Hazel. *The Invasion from Mars*. Princeton University Press, 1940.
3. Beville, Hugh. *The Social Stratification of the Radio Audience*. No. 1 of Studies in the Social Psychology of Radio, Office of Radio Research, November, 1939.
4. Lazarsfeld, Paul F., Berelson, Bernard, and Gaudet, Hazel. *The People's Choice*. Duell, Sloan and Pearce, 1944.
5. Meyrowitz, Alvin, and Fiske, Marjorie. The Relative Preference of Low Income Groups for Small Stations. *Journal of Applied Psychology*, 23, No. 1 (February, 1939).
6. Fiske, Marjorie. *A Sampling of Public Attitudes Toward the Proposed Subscription Radio Plan*. Office of Radio Research, December, 1944.
7. Curtis, Alberta. *Listeners Appraise a College Station*. Federal Radio Education Committee Publication, 1940.
8. Robinson, William S. Radio Comes to the Farmer in *Radio Research—1941*. Duell, Sloan and Pearce, 1941.
9. Fiske, Marjorie. How Consumers React to Radio Advertising by Retailers. *Sales Management*, 35, No. 8.
10. Hallonquist, T., and Suchman, E. A. Listening to the Listener, *Radio Research, 1942-3*. Duell, Sloan and Pearce, 1944.
11. Herzog, Herta. On Borrowed Experience. *Studies in Philosophy and Social Science*, 9, No. 1 (May, 1941).
12. Herzog, Herta. What Do We Really Know About Daytime Serial Listeners? in *Radio Research, 1942-3*. Duell, Sloan and Pearce, 1944.
13. Arnheim, Rudolph. The World of the Daytime Serial, *ibid*.
14. Fleiss, Marjorie. The Panel as an Aid in Measuring the Effects of Advertising. *Journal of Applied Psychology*.
Lazarsfeld, Paul F., and Fiske, Marjorie. The Panel as a New Tool for Social Research. *Public Opinion Quarterly*, 2, No. 4 (October 1938).
Smith, C. H. The CBS Forecast Panel in *Radio Research 1942-3*. Duell, Sloan and Pearce, New York, 1940.
15. Lazarsfeld, Paul F. Audience Building in Educational Broadcasting. *Journal of Educational Sociology*, 14, No. 9 (May, 1941).
16. McCandless, Boyd. A Study of Non-Listeners, in *Radio Research 1942-3*. Duell, Sloan and Pearce, 1944.

Chapter XI

THE NEW COOPERATIVE ANALYSIS OF BROADCASTING

A. W. LEHMAN AND GEORGE H. ALLEN
Cooperative Analysis of Broadcasting

THE Cooperative Analysis of Broadcasting, the only co-operative, non-profit organization measuring radio listening, expanded in 1944 both its field of research and the information in its reports. The expansion, to what is virtually a new C.A.B., is founded on sixteen uninterrupted years of experience in investigating and reporting the size of radio audiences.

By this expansion, the C.A.B.'s Board of Governors has directed its sights toward postwar perfection of its program ratings. The expansion, according to the board's announcement, is "an essential . . . stride toward its eventual goal of a rating service representing a true cross-section of all American homes, urban and rural, telephone and non-telephone." Thus the C.A.B. has set down for its members the forthcoming pattern of research in radio listening.

Years of experimentation have led the C.A.B. to rely more and more on the coincidental technique in its investigations. Many experiments with various techniques have resulted in the open-minded realization that each method carries its own limitation. For example, C.A.B. ratings for late-at-night programs in the Eastern Time Zone are the result of the combined use of the day-part recall and coincidental techniques. All other ratings are produced by coincidental telephone interviews, producing an average audience rating.

At present, the dialings on which the C.A.B. bases its ratings exceed 6,300,000 a year. The sampling is nearly three times the coincidental quota for C.A.B. reports in 1943 and is 60 times as large as the 100,000 used in its early days.

Interviews are distributed among 81 cities of 50,000 or more in population in contrast with the previous field of 33 cities. The sample is spread so that the number of calls in cities of each size group is proportioned in accordance with the population contained in the cities within that population group.

152 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

Thus, for example, 16 per cent of the sample is collected in cities of 50,000 to 100,000 in keeping with the location of 16 per cent of the urban population in cities of that size. Incidentally, none of the 33-city sample was allocated to cities in that group.

The proration of the sample according to city size is shown in the following table, with an interesting comparison of the distribution of the 33-city sample:

City Size	Population Per Cent	Former Sample	New Sample	Pts. Chg. from 33 to 81 Cities Per Cent
		33 Cities Per Cent	81 Cities Per Cent	
Over 1,000,000	35	20	36	+16
500,000 to 1,000,000	14	24	14	-10
100,000 to 500,000	35	56	34	-22
50,000 to 100,000	16	0	16	+16
Total	100	100	100	

By thus giving each city-size group its true representation, the C.A.B. has achieved what it considers an adequate representation of the United States cities of 50,000 population and up. As the table indicates, the sample in cities of 1,000,000 or more had to be increased 16 percentage points from its former proportions to produce a representative base. In contrast, the sample in the intermediate groups of cities (100,000 to 500,000 and 500,000 to 1,000,000) was reduced percentagewise, although, as a result of the over-all increase in dialings, the sample is numerically larger.

A similar adjustment occurred in the distribution of the new sample in nine geographical areas defined by the United States Census. The C.A.B. formerly used four geographical areas as a basis for its sectional rating reports. The new distribution is shown in the following table where it is contrasted with the allocation of the former sample:

	Population Per Cent	Former Sample	New Sample	Pts. Chg. from 33 to 81 Cities Per Cent
		33 Cities Per Cent	81 Cities Per Cent	
New England	7.4	8.8	7.4	- 1.4
Middle Atlantic	32.6	20.6	33.6	+13.0
East North Central	24.2	14.7	23.9	+ 9.2
West North Central	7.2	14.7	7.4	- 7.3
South Atlantic	8.7	8.8	8.2	- 0.6
East South Central	3.4	5.9	3.0	- 2.9
West South Central	5.7	11.8	6.0	- 5.8
Mountain	1.3	0	1.5	+ 1.5
Pacific	9.5	14.7	9.0	- 5.7
Total	100.0	100.0	100.0	

As with the redistribution of the sample among cities, the new apportionment among *geographic areas* testifies to the adequate and representative quality of the new C.A.B. base. By the same token, of course, it reveals the inadequate allocation of any base that does not properly represent the distribution of population, by sections of the country, and by city-size groups, etc. The C.A.B. continues to report its ratings by the 33-city sample as well as those recorded from the massive base in 81 cities. The ratings reckoned from the lesser base have been retained on a somewhat temporary footing to provide comparative continuity with the ratings recorded prior to the expansion in April, 1944.

The present cross section is a foundation on which the C.A.B. can expand its sample without distortion. The ultimate goal, of course, is to include non-telephone homes in all urban areas and to extend coverage to rural areas and urban communities of less than 50,000. In its current reports, the C.A.B. specifically suggests that the 81-city ratings represent only the listenership registered in the telephone-home population of the C.A.B. cities. The C.A.B. reports carry even more firm warnings against the recurring temptation to project audience ratings over the nation's total number of radio homes.

Tripling the dialings has resulted, of course, in a similar increase in the base for rating individual programs. The nighttime sample for each half hour of broadcasting has been enlarged 159 per cent while the quarter-hour daytime unit figured as five broadcasts a week on Monday through Friday, is now measured by a base 158 per cent larger than in 1943.

Ratings for each half hour of nighttime broadcasting are now computed from a foundation of 4,020 calls in 81 cities. Programs of lesser geographical scope are measured with a base proportionately smaller, in proportion to the program's coverage. Less extensive broadcastings are rated thus in the actual terms of their potential audiences.

Because the C.A.B. now covers one, two and three network as well as four network cities, it was decided to include in the city schedules on which a program's rating is based, the cities covered "effectively" from without. "Effective coverage" is reckoned separately for day and night broadcasting at regular intervals throughout the year and is the first time in network radio that a network's coverage from without a city has been measured. The procedure first measures the average listening to a given network in a city where it does not have an outlet. This average is then compared with its average in all other cities where the network has stations. If the average in the non-station city is at least on a par with the average in station cities, the network is credited with effective coverage. Thus the base for rating the networks offerings is extended to such nonstation areas.

The nighttime C.A.B. sample of 4,020 calls per half hour represents, of

course, the dialings for that time span in two calendar days in the two interviewing weeks covered. A 15-minute nighttime broadcast running five times a week has as its 81-city sample a total of 10,050 calls. The hour program, produced once a week, has simply twice as many dialings in its base as the half-hour program. All ratings are based on the results of two full weeks of interviewing.

The daytime base is exactly half the size of the nighttime sample. Thus the five-a-week programs of 15 minutes' duration are measured with an 81-city sample of 5,025 calls. The former 33-city base, incidentally, was 1,950. Daytime programs broadcast on Saturday and Sunday, however, are rated from a base equal to the nighttime sample.

The C.A.B. uses the coincidental method of investigation. Interviewers ask what program and what station are being listened to at the time the telephone rang. People who answer "Don't know" are included in the sample. Homes dialed and found "Busy" on the phone are excluded from the sample. The investigation periods run continuously through the day for programs originating from 9:00 A.M. to 10:30 P.M. The sampling is collected at random from the telephone homes in the 81 cities forming the C.A.B. national cross section.

Programs broadcast in the Eastern Time Zone between 10:30 and 11:30 P.M., however, are rated by the equivalent of their coincidental audience. The first step in reckoning the late-hour coincidental for Eastern seaboard broadcasts is to record their coincidental ratings in the Central Time Zone. The same programs are then rated by the day-part technique the next morning in both the Eastern and Central Time Zones. Listeners are asked to recall the programs they heard the previous night in the hour ending with 11:30 P.M. The Central coincidental rating on each program is divided by the Central day-part rating on each program. The resulting factor, the quotient, is used to adjust the Eastern day-part rating. The formula thus employed gives a product that represents the estimated coincidental rating of the late-night programs in the East.

The new 81-city sample, the largest cross section of the nation ever used for a continuous survey of network audiences, has also afforded many corollary benefits to the C.A.B. member.

Ratings three times a year are now released on network programs for each of four city-size groups—for cities 50,000 to 100,000; 100,000 to 500,000; 500,000 to 1,000,000; 1,000,000 and over in population. Ratings are also released three times a year for the first time, based on seven United States geographical areas as contrasted to the normal four geographic areas to which network radio has become accustomed.

This added and intensive C.A.B. coverage of more cities, and more types

of cities, has resulted in the C.A.B.'s being able to release to members twice-a-month ratings on limited network and regional network programs, ratings by the company sales territories of national advertisers, and ratings on nationally broadcast *live* or transcribed "spot" radio programs. All of these innovations have now become a standard part of the C.A.B.'s national research service.

Another decided innovation has been the publication by the C.A.B. of six reports a year on the radio audience in each of the 81 cities in which it surveys regularly. These individual city reports reflect the audiences of approximately 300 radio stations, and evaluate the average program popularity of these stations six times throughout the seasons of the year. A particular innovation to the local radio scene is the strict comparability of all of these local data by time of day, date of interviewing, interviewing method, etc.

Of special interest during 1944 has been the doubling of the participation in the C.A.B. of advertisers, agencies, and other organizations interested in network radio. This 100 per cent increase in the number of participating concerns marks a healthy step forward for all advertising toward the goal of a more efficient postwar marketing system. Network radio in particular seems to be preparing itself for a more intensive evaluation of its audiences postwar as well as now.

Chapter XII

THE COINCIDENTAL METHOD OF MEASURING RADIO AUDIENCE SIZE

C. E. HOOPER
C. E. Hooper, Inc.

AMERICAN advertisers spent an estimated \$228,764,705 in the purchase of radio time in 1943. [1] One company alone spent \$10,825,222 for this service. [2] The purchase of radio time is guided largely by audience measurements obtained by use of the "coincidental" method. Probably no other act based on sampling human behavior carries so heavy a burden of financial responsibility.

The coincidental method gets its name from the fact that all interviews used in determining the characteristics of the listening audience are obtained while the program is being broadcast. It was not the first method for obtaining program audience measurements to make its appearance. Both the unaided recall (which was changed to day-part) and the aided recall, or printed roster, preceded it. It is, nonetheless, now the most widely used of the methods for obtaining audience measurements, both for individual program audiences and for station audiences. Its wide acceptance and use makes its examination the more necessary.

The coincidental method, developed for most part by C. E. Hooper, Inc., has recently (1944) been adopted by the Cooperative Analysis of Broadcasting, but because the C.A.B. experience is so recent, the method will be discussed here as it has been used by C. E. Hooper, Inc., continuously since 1934. The method has many outstanding advantages which appeal to research men. First, being an interview, it gets data from direct contact with the respondent, the only person qualified to report the fact of listening or not listening. Secondly, it is a short simple interview with a minimum of influence on the respondent and requires a minimum of effort on the part of the respondent. Third, it has a fundamental characteristic found in no other

method: it measures listening to the radio during the time each program is on the air—the only time at which the program's audience exists. Fourth, the interview may be conducted readily by telephone or by personal interview. Currently most of the interviewing is done by telephone.

The coincidental questions used by C. E. Hooper, Inc., are the following:

1. WERE YOU LISTENING TO YOUR RADIO JUST NOW?
2. TO WHAT PROGRAM WERE YOU LISTENING, PLEASE?
3. WHAT STATION PLEASE?
4. WHAT IS ADVERTISED?
5. HOW MANY MEN, WOMEN AND CHILDREN IN YOUR HOME WERE ACTUALLY LISTENING?

When a Hooper interviewer attempts to contact a home by telephone, any one of a number of conditions may be encountered.

1. The operator may report that the telephone "has been disconnected."
2. There may be no answer.
3. A "line busy" signal may be obtained.
4. The ring may be answered by a man, a woman, or a child.
5. The respondent may refuse to be interviewed.
6. The respondent may report that he is (a) not listening to the radio, (b) listening to the radio.
7. If he is listening to the radio, he may report (a) the program name or the name of identifiable talent or (b) that he does not know the name.
8. He may report (a) that he does or (b) that he does not know to what station he is listening.
9. He may report (a) that he does or (b) that he does not know who sponsors the program to which he is listening.
10. He will usually report how many men, women, and children are listening.

From the above responses a wide range of useful figures are developed.

1. *Disconnects.* All "disconnects" are eliminated from the sample and treated as though the call had never been made. These data are used only to yield an index of family movement within the given locality.

2. *Don't answer.* If there is no answer after the telephone has rung six times, it is assumed that there is no one "at home and awake." Such homes are classed as "nonlistening" homes. By subtracting the per cent of "not at home and awake" from the total sample, the per cent of "at home and awake" is obtained. This is an important segment of the population to the radio advertiser and is called the "available audience." It gains its importance from the fact that the advertiser can recruit his listening audience only from the available audience. The smaller the available audience becomes, other

things being equal, the smaller is his opportunity for recruiting a listening audience of given magnitude.

3. *Line busy.* A "line busy" signal also conveys important information. Such a signal is assumed in the Hooper operation to show that someone is "at home and awake." While we do not discover whether or not these homes are listening, it is further assumed that listening in these "at home and awake" families is distributed as it is in other "at home and awake" families. That is, "lines busy" are prorated to "listening" and "nonlistening" instead of adopting the more drastic alternative of dropping them out of the sample.

Let us consider the basic assumption made—that a "line busy" signal shows that someone is "at home and awake." First, it is clear that complete justification of this assumption would require that all residence main telephones be private lines. Even before the war, this was not the case. During the course of the war, the per cent of party lines increased somewhat. A "line busy" signal on a party line shows only that some one of the parties is talking and not necessarily the party called. Since "lines busy" may constitute 10 per cent of the sample, it is sometimes thought that we introduce an important source of error as the number of party lines increases.

This problem has recently been considered in detail in an article which appeared in *Broadcasting Magazine*. This article points out the following: We know from our available audience figure that, depending on the time of the day and season of the year, someone is "at home and awake" in from 70 to 85 per cent of the homes. We know this from past experience before we start calling. It would be expected to be as true of a sample of homes in which a "line busy" signal is subsequently obtained as it would for a sample where no "line busy" signal is subsequently obtained. That is to say, regardless of intelligence conveyed by the signal, we know from our available audience figure that in the evening in September, for example, someone is "at home and awake" in 75 per cent of these homes. With this knowledge we would not be justified in dropping more than 25 per cent of the "lines busy."

But the probability that someone is "at home and awake" in such homes does not stop there. Suppose, for example, that 50 per cent of the residence main telephones in a locality are private wires and 50 per cent are party lines. A "line busy" signal on the former shows that someone is "at home and awake" in all of them. It is, therefore, only the remaining 50 per cent party lines to which we should apply the 75 per cent probability, which is to say 37.5 per cent. This, added to the 50 per cent, indicates that under these circumstances 87.5 per cent of the "lines busy" would logically be allocated to "at home and awake."

But even if we adopt 75 per cent, a most conservative figure, and, as in

Table I, calculate program audience sizes, first (Column 1), by prorating 75 per cent of the "lines busy" and, second (Column 2), by prorating all the "lines busy" we obtain the figures shown for a high, a medium and a low rating evening program and for a high and a low rating daytime program.

TABLE 1
EFFECT OF DIFFERENT TREATMENTS OF "LINES BUSY"
ON 32-CITY "NETWORK HOOPERATINGS"

	Column 1 Prorating 75% of "Lines Busy"	Column 2 Prorating 100% of "Lines Busy"	Column 3 Eliminating All "Lines Busy"
<i>Evening Programs</i>			
Fibber McGee & Molly	30.2	30.3	29.8
Truth or Consequences	11.1	12.2	11.9
Spotlight Bands	1.7	1.7	1.6
<i>Daytime Programs</i>			
Right to Happiness	7.4	7.5	7.3
Vic and Sade	3.0	3.0	2.8

It appears from the results in Table 1 that, in comparison with the most conservative conditions, prorating of all "lines busy" does no injury to the calculation of audience size.

If, however, we calculate the audience size omitting the "lines busy" entirely we obtain results shown in Column 3 of Table 1. The difference between these and either Column 1 or 2 becomes slightly more marked. In this case the audience size appears to be uniformly underestimated.

While some sampling organizations eliminate "lines busy" from their sample, it is difficult to perceive grounds for doing so other than the fact that it simplifies the calculation and reduces costs somewhat.

4. *Phone answered.* When the phone is answered, we know whether the respondent is a man, woman, or child. Question 5 above determines age and sex of actual listeners, which varies widely by programs, time of day, day of week, season of year. The fact of answering tells us someone is "at home and awake," and is, therefore, available for listening.

The daytime and evening trends of available audience from January, 1942, through October, 1944, are shown in Charts I and II respectively. Examination of the available audience trend lines shows that they tend to be highest in January and February and lowest in July and August. The range of variation in the evening over this three-year period has been from 70.2 in July, 1944, to 83.2 in January, 1943. The range of variation for daytime during the same period has been from 68.9 in August, 1944, to 75.6 in February, 1943. It will be noted that available audience in both daytime and evening has decreased as the size of our armed forces and industrial employment have increased.

CHART I
Hooper National Trends
AVAILABLE AUDIENCE INDEX -- DAYTIME
 JANUARY, 1942 - OCTOBER, 1944

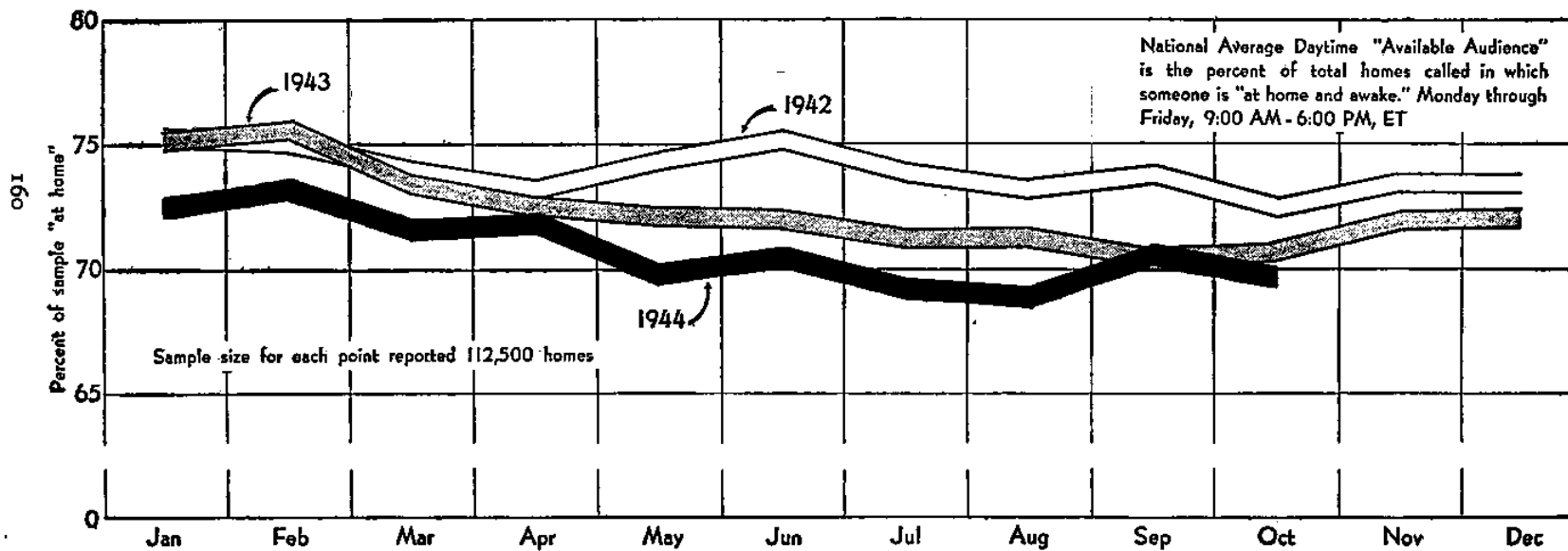
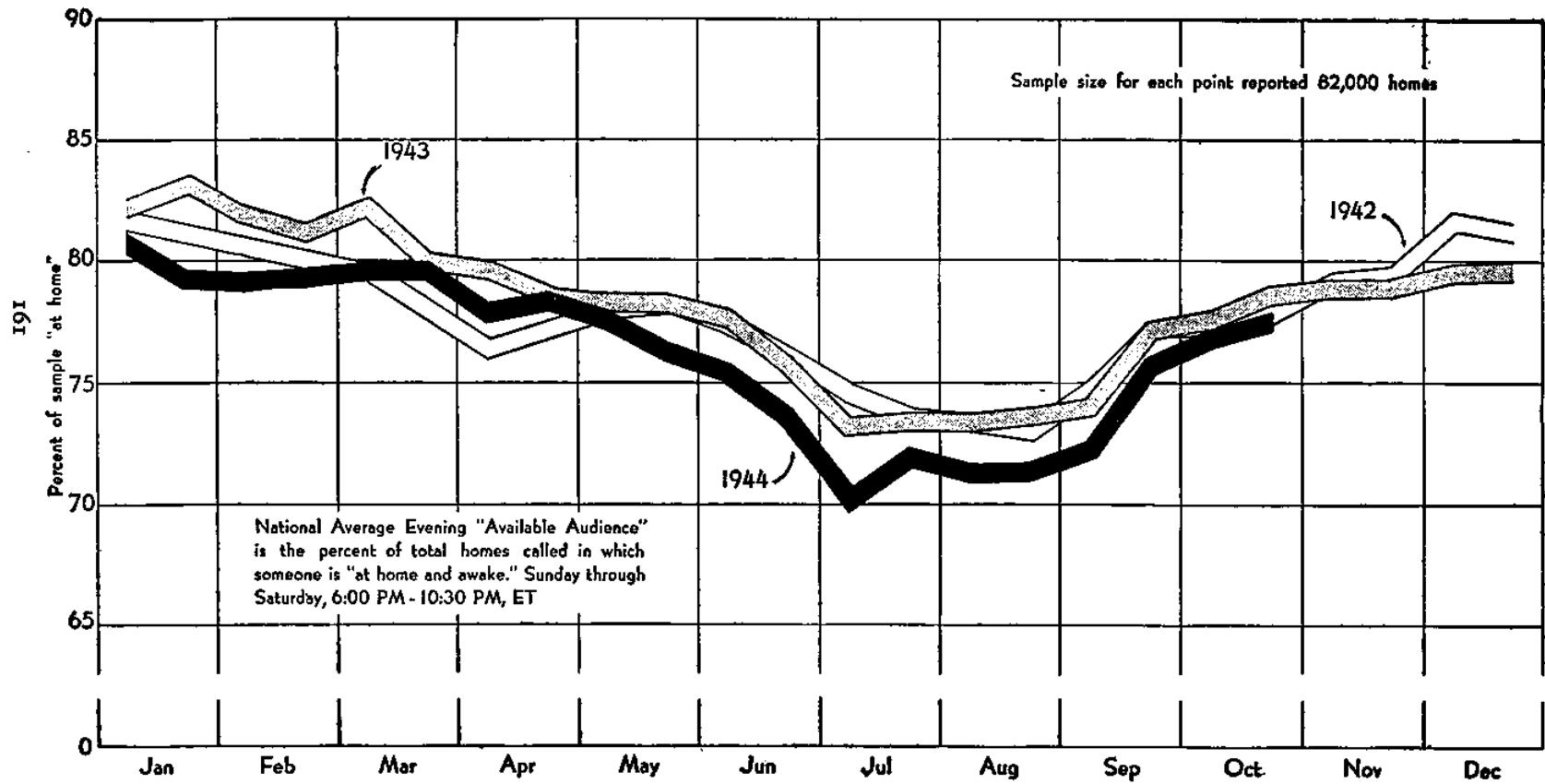


CHART II
Hooper National Trends
AVAILABLE AUDIENCE INDEX -- EVENINGS
 JANUARY, 1942 - OCTOBER, 1944



5. *Refusal to be interviewed.* If the respondent refuses to answer because she is busy with the baby or other activity, we know that there is someone "at home and awake." Therefore, the refusal becomes a part of the base for calculating available audience.

Every method of radio audience measurement is dependent upon the co-operation of the public including both listener and nonlistener. It becomes necessary for the measurer so to conduct himself as to preserve a public free from hostility or expressed irritation. The public is extremely co-operative except in localities where it has been abused by overinterviewing. Hostility can be expected of interviewing if conducted before 8:00 A.M. or after 10:00 to 10:30 P.M. in most communities.

6. *Listening or not listening.* If the respondent reports that he is not listening to the radio, the call also becomes part of the base for calculating the available audience. If, on the other hand, he reports that he is listening, the call may be used not only in calculating the size of the available audience but also in calculating the per cent of homes listening to some program at the given time. This figure, showing average listening during a given time period, is called "sets-in-use." The daytime sets-in-use trend from January, 1942, through October, 1944, is shown in Chart III, while that for evening hours is shown in Chart IV.

It will be seen immediately that there is far more listening in the evening than in the daytime. It will also be noted that sets-in-use undergo much more violent seasonal changes in the evening than in the daytime. In part, this results from the fact that many high-rating evening programs are not broadcast in the summer but are replaced by substitutes. Daytime programming, on the other hand, remains much the same from winter to summer.

7-8. *Program-station identification.* If the respondent is listening to the radio, he may or may not be able to report the name of the program or identifiable talent. If he cannot, he may report the call letters or the dial setting of the station to which he is listening. Or he may know the name of the program's sponsor.

All these pieces of intelligence serve to identify the program to which he is listening and to form the base for calculating the size of the audience listening to specific programs or to specific stations. This figure shows *the per cent of total homes called which were listening to a specific program and is called the "Hooperating."* As was pointed out above, "lines busy" and "refusals" are a part of the "at home and awake" segment of the sample, but since we have no listening reports from them, it is assumed that their listening is distributed in the same manner as is that of the much larger part of the "at home and awake" group. That is to say, they are prorated over all the programs reported.

CHART III
Hooper National Trends
SETS-IN-USE INDEX -- DAYTIME
JANUARY, 1942 - OCTOBER, 1944

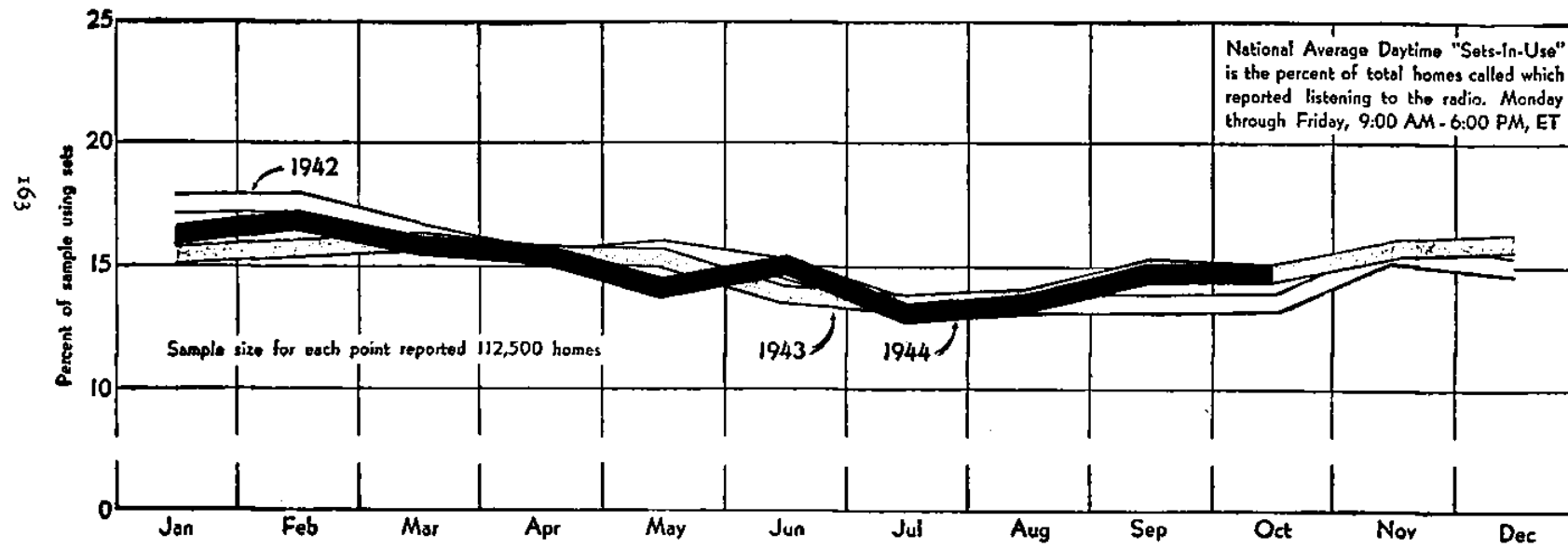
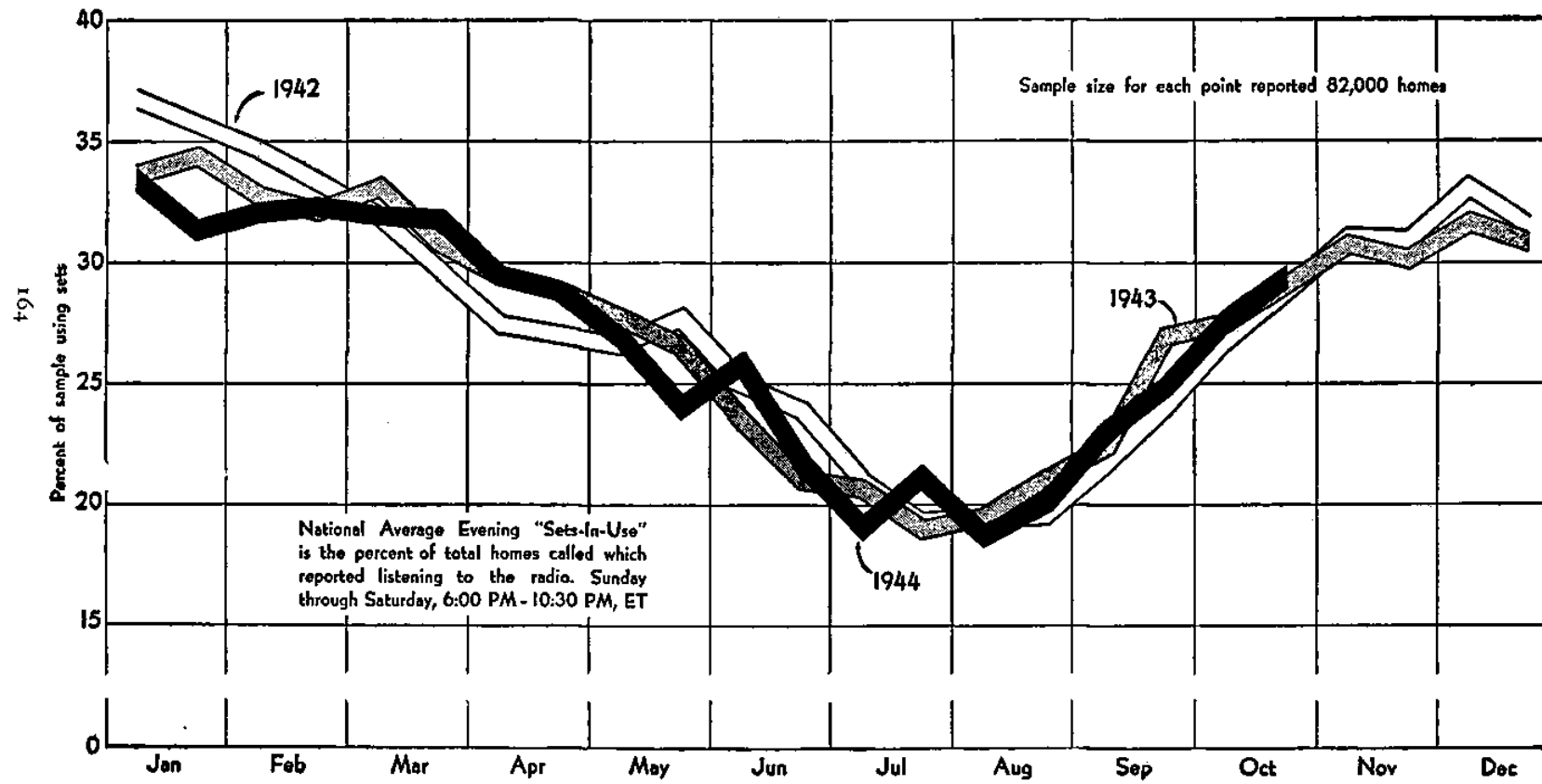


CHART IV
Hooper National Trends
SETS-IN-USE INDEX -- EVENINGS
 JANUARY, 1942 - OCTOBER, 1944



The average 32-City "Network Hooperatings" trends for daytime and evening are shown in Charts V and VI, respectively. Like sets-in-use, average ratings tend to be higher for evening than for daytime programs and seasonal fluctuations are greater for the evening. It is interesting to note that while listening to the radio (sets-in-use) was about as great in 1944 as in previous years, ratings were lower in 1944. This results from the fact that many new programs were added in 1944 and new programs start with low ratings, pulling down the average.

Identification of the program or station listened to also provides another measure. This per cent is called "share-of-audience." This is quite different from the audience size figure, or rating, which is based on total homes called (excluding disconnects). Similarly, the use of the figures is quite different. For example, a program obtaining a rating of 8.0 in the daytime would have a 50 per cent share-of-audience if 16 per cent of the sets are in use; a program obtaining the same rating of 8.0 on Sunday evening when sets-in-use are 40 per cent would obtain a 20 per cent share-of-audience. Further, a program might get a rating of 10.0 in February with a 35 per cent share-of-audience. In August the rating might have fallen to 6.0, but the share-of-audience might still be 35 per cent. Share-of-audience is a measure of a program's "pull" against its coincidental competition, and its usefulness as such is enhanced by its freedom from influences of change in the size of the total listening audience from season to season.

9. *Sponsor identification.* When the respondent reports the name of the advertiser or the name of the product advertised or that he "does not know," these data furnish the base for calculating the Sponsor Identification Index. This figure shows the per cent of those who were found listening to the program who:

1. Identified the advertiser or the major product advertised.
2. Partially identified the sponsor.
3. Misidentified the sponsor.
4. Did not know the sponsor.

10. *Who is listening.* When the respondent reports on the number of men, women, and children listening to the program he furnishes data for analysis of the composition of the audience in the following ways:

1. Average number of persons listening per listening set.
2. Average number of women listening per listening set.
3. Average number of men listening per listening set.
4. Average number of children listening per listening set.
5. Per cent of all homes sampled with *men, women, and children* listening.
6. Per cent of all homes sampled with *men and women* listening.

CHART V
Hooper National Trends
AVERAGE RATING INDEX -- DAYTIME
 JANUARY, 1942 - OCTOBER, 1944

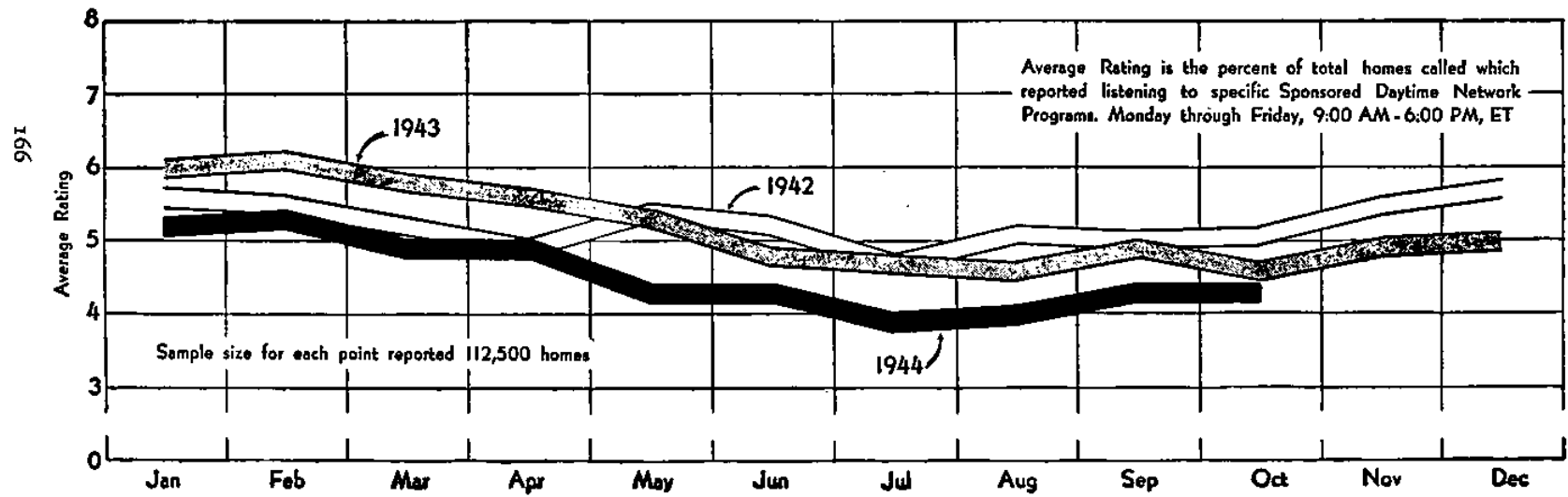
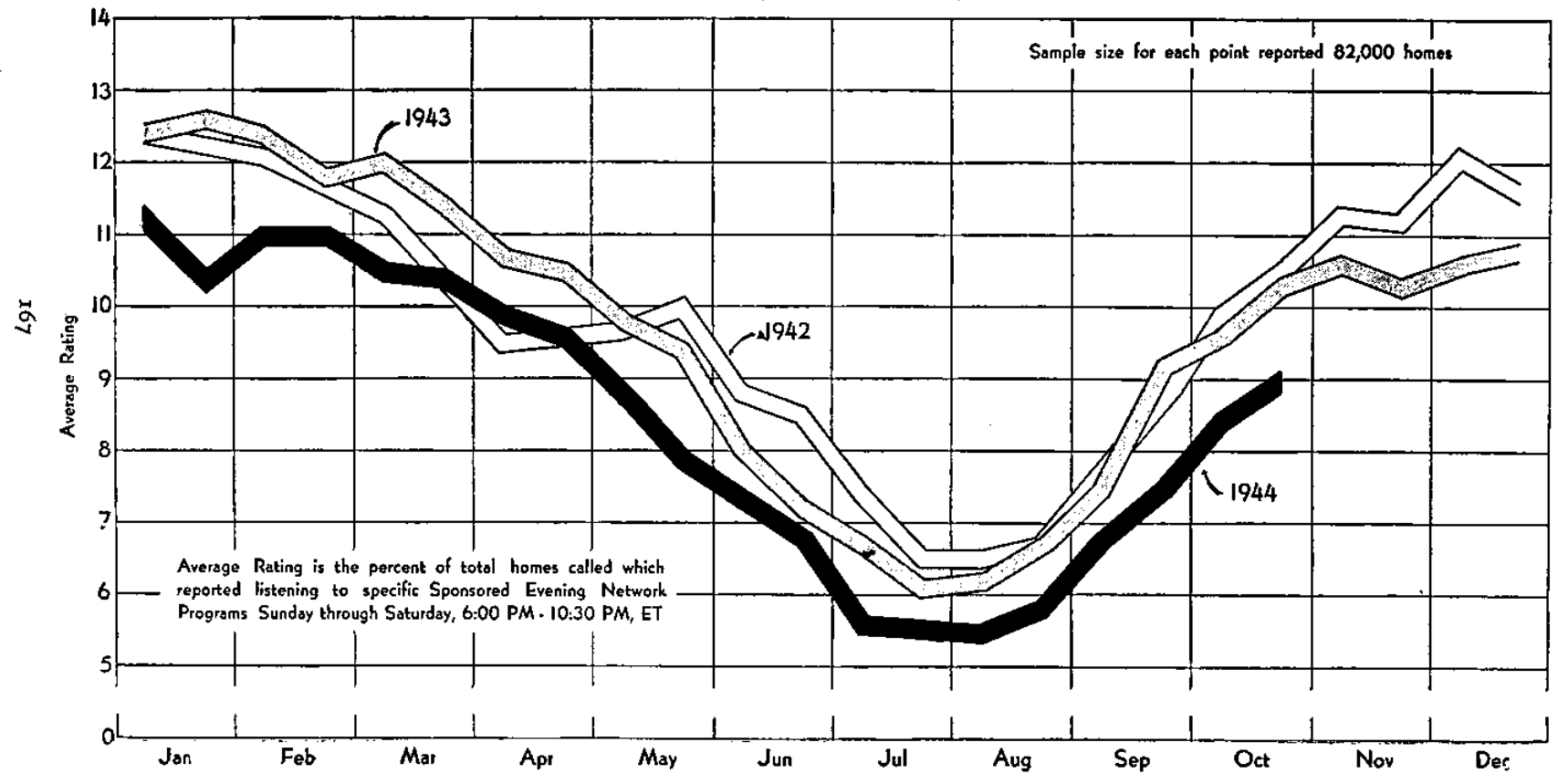


CHART VI
Hooper National Trends
AVERAGE RATING INDEX -- EVENINGS
 JANUARY, 1942 - OCTOBER, 1944



7. Per cent of all homes sampled with *men and children* listening.
8. Per cent of all homes sampled with *women and children* listening.
9. Per cent of all homes sampled with *men only* listening.
10. Per cent of all homes sampled with *women only* listening.
11. Per cent of all homes sampled with *children only* listening.

From these data it is possible for the advertiser to determine with some accuracy whether or not his program is reaching the audience who form the potential market for his product.

CHART VII
**AGE AND SEX GROUPS IN
 HOMES LISTENING TO "LONE RANGER"**
 MAY - AUGUST, 1943
 100% = 363 homes listening to "Lone Ranger"

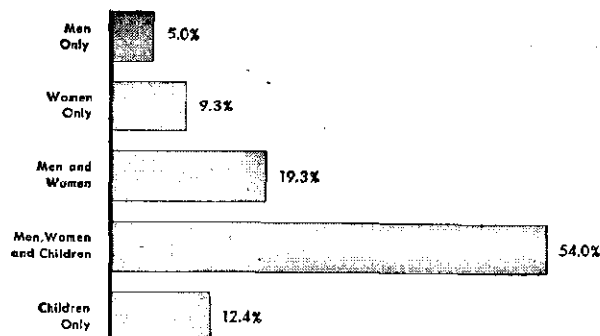


Chart VII shows the composition of "Lone Ranger" audience. This program is designed primarily to reach young folks, and it does. But it also reaches adults very well.

Currently, women predominate in audiences in the evening as well as in the daytime, though to a lesser degree.

USE OF COINCIDENTAL DATA

Numerous audience-size measurements are required in the radio field, each of which has a use quite distinct from that of any other and each of which must, therefore, be obtained from a different sample. Four are described below.

1. 32-City "Network Hooperatings." The first type of coincidental telephone radio audience measurement to be produced by C. E. Hooper, Inc., was

what has now come to be known as the "32-City Network Hooperating." The function of this figure is to show for a network program the "pulling power" or attraction value of the program under conditions of wide but uniform freedom of choice between programs.

In order to obtain such a measure of network program attraction value, it is obvious that the sample can be taken in only those places where maximum freedom of choice can be exercised. For example, if only one network can be heard in a given locality, the only choice the listener can make is to listen to the network show or not to. If two networks can be heard in the locality, the choice is greater. But not until we reach cities in which all four networks are heard with ease do we have conditions of maximum freedom of choice between network programs. To say it another way, in these cities it is possible to eliminate to a maximum degree variations in network competition. They are, in short, areas of maximum network competition.

Another factor that must be controlled in measuring the audience pulling power of a program is the variable of differences in "coverage" between programs. Obviously, if a program is carried on 135 stations of a network it would be expected to have a bigger audience than if it were carried on 32 stations of the same network; but that would involve no change in the pulling power of the programs. It follows that the measurements of attraction value, or audience pulling power, must be made under conditions where differences in program coverage also are controlled.

As a result of intensive and extensive studies of listening areas in the United States, C. E. Hooper, Inc., selected 32 cities, in each of which there is a local outlet for each of the four major networks, as satisfying the standards required for developing such ratings. Within the 5-cent telephone zones, those cities are free from variations in program coverage because the network outlet can be heard everywhere the sampling is done. In every respect they present as uniform conditions of network competition as it is possible to obtain. They are "areas of uniform network competition," but it should be borne in mind that they are also the "areas of maximum network competition."

Under conditions where the variables of coverage and competition are controlled with as high degree of precision (as they are in the C. E. Hooper, Inc., 32 four-network cities), and only under such conditions is it possible to obtain ratings which are measurements of network-program audience pulling power, or attraction value. 32-City Network Hooperatings are such a measure.

Reports containing coincidental data as obtained under these standard conditions are called "Network Hooperating Reports." They are published twice monthly for evening programs. Subscribers to this network program

rating service are: the networks, national advertisers, advertising agencies, radio talent, radio station representatives, and the government.

2. *City Hooperatings*. In the early days of radio audience measurements the interest was almost entirely in sponsored network program ratings. Radio station popularity was measured largely by incoming mail. However, it soon became apparent that radio listening habits were, to a large extent, local in character rather than national. For example, Denver and Salt Lake City are two areas of "equal network opportunity" in the Mountain region. However, Station KSL in Salt Lake City may deliver a local rating of 40.0 to "Lux Radio Theater," while Station KLZ in Denver may deliver only a 22.0. Numerous local factors, even in the same geographical area, play an important part in determining the size of the audience attracted by a program in each locality. The interest in local measurements was further facilitated by the rise of spot broadcasting and by the needs of the local stations which have no network affiliations.

To meet this demand, C. E. Hooper, Inc., developed in 1940 a "Continuing Measurement of Radio Listening" service for individual cities. In this service the local stations of each city studied are furnished with (1) comparative day-part indexes to station listening for individual stations for the morning, afternoon and evening hours, and (2) programs and time-period ratings.

The utility of these studies was demonstrated by the fact that the local stations in 70 large cities adopted them. Another 200 cities which are too small for continuous study subscribe for intermittent day-part Station Listening Indexes.

3. *Area Hooperatings*. The inherent limitation of the city reports lies in the fact that they measure mainly listening in the city proper. Station signals go beyond city limits. Area Hooperatings have, therefore, been developed to determine program audiences in the total area served by the station—this area being defined by the station. [2]

The telephone coincidental which provides the swiftest and most economical audience measurements in the big cities has, of necessity, been supplanted in the area Hooperatings by a different method. After exhaustive research we have chosen the "diary" or family listening record. (The use of this technique, for this specific purpose, was suggested by us as far back as 1943 in Chapter IX of a book, *Radio Audience Measurement*, which every station manager has on his desk.)

The audience measurements provided by the diary correlate very closely but at a higher level with the findings of the telephone coincidental method. A major reason for the higher level results from part-program listeners in any quarter-hour periods being equally weighted with listeners during the entire quarter-hour. On the other hand, the diary technique develops some

interesting by-products of its own. These include "flow of audience" ("from where" and "to where"), "duplication" of listeners of different programs and "frequency" of listening to individual programs. Most of the inflation present in diary records as a result of "tuning" unaccompanied by "listening" has been wrung out as a result of refinements we have introduced into the technique.

In area Hooperatings provided for stations located in cities which we service with city reports, the diary results of the area survey are converted to the coincidental base. This permits direct comparison between city and total area audiences. In other communities where the station city-size is inadequate to support city Hooperatings, only the diary results are published. These reports warn against the misinterpretation of diary findings as audience measurements comparable with coincidental measurements.

4. *National Hooperatings*. This service is designed to give advertisers a measure of the truly national audience to specific network programs. It is the final, long step in radio audience measurement, and it is our expectation that this service can be put into operation during 1946 on a practical, economical basis.

National Hooperatings would be based on "diaries" in 7,500 or more radio homes, carefully selected to represent, in proper balance, network service, population distribution, etc. It is essentially an extension of the technique already thoroughly established in connection with area Hooperatings. Diaries would be installed for a one-week period for each report, of which there would be two, or at most three, per year. Interim national audience estimates could be provided through correlation with "network Hooperatings" in 32 cities, a service which would, of course, be continued.

This plan for a national Hooperating service, immediately recognized as a new measurement of actual program listening on a truly national basis, has already been discussed with the research directors of the four networks.

Those who are primarily interested in the technical problems of radio audience measurement, such as comparative validity of different methods, the adequacy of techniques of sampling by telephone and by personal interview, the variables encountered in the broadcast situation, the reliability of sample size and similar scientific matters, can find them discussed at length in *Radio Audience Measurement*, Stephen Daye, New York, 1944.

The writer has attempted in this chapter to show only what the coincidental method is and the purposes for which it is most commonly used.

REFERENCES

1. Compiled from F.C.C. Analysis of Economic Status of Broadcasting. *Broadcasting Magazine*, Oct. 16, 1944, p. 14.
2. Description Area and National Hooperatings from *Broadcasting*, March 18, 1946.

Chapter XIII

THE PERSONAL INTERVIEW-ROSTER METHOD OF RADIO MEASUREMENT AND ITS APPLICATION

SYDNEY ROSLOW AND NELLE KELLY
The Pulse, Inc.

IN 1941, after a series of successful experiments with personal interviewing by means of an aided recall roster method, the Pulse of New York,¹ under the direction of Dr. Sydney Roslow, entered the field of radio audience measurement with its monthly survey of listening in the metropolitan New York area. At that time there were already in existence other means of audience measurement, namely, the telephone recall and telephone coincidental. The personal door-to-door interview method has certain advantages over both of these.

It reaches both telephone and nontelephone homes. In any measurement of public habits and attitudes, an adequate sampling of all elements of the population should be obtained, in order to achieve accurate results. The telephone survey reaches only that 40 per cent of the population (in New York City) who own telephones, and these telephone homes are not representative of the total population. Actual breakdowns of audiences to many programs show significant differences in the listening habits of the telephone and nontelephone groups (see Tables 1 and 2).

"Class" programs, such as symphonic music, forums, etc., are given a higher rating by the telephone method in relation to sets-in-use than they actually deserve, while programs like the Make Believe Ballroom, consisting of popular recordings, show up to disadvantage because the bulk of their audience comes from the lower income nontelephone portion of the population.

Better rapport between respondent and interviewer can be established by means of a personal interview, and more information can be obtained for

¹ Now The Pulse, Inc.

TABLE 1
COMPARISON OF RATINGS OBTAINED BY TELEPHONE COINCIDENTAL
AND
PERSONAL INTERVIEW-ROSTER TECHNIQUES

(March-April averages, 1944)

	TELEPHONE COINCIDENTAL		PULSE ROSTER	
	Ratings %	Sets-in-Use %	Ratings %	Sets-in-Use %
MAKE BELIEVE BALLROOM WNEW, Mon.-Fri. 6-6:15 PM	3.0	19.2	5.5	23.9
TOWN MEETING OF AIR WJZ, Thurs. 8:30-9:30 PM	3.2	39.8	3.6	51.4
JOHN CHARLES THOMAS WABC, Sunday, 2:30-3:00 PM	5.8	20.0	4.9	20.9
GOOD WILL HOUR WMCA & WOR, Sun 10:15-10:30 PM	4.9	38.8	6.7	37.5
CONTENTED HOUR WEAF, Monday 10-10:30 PM	6.8	34.2	5.1	32.1

purposes of classification. Telephone interviewing is restricted to the simplest and briefest form of question because of the need of making a given quota of calls within a limited time period. Such questions as age, socioeconomic status, education, and nationality cannot be obtained over the telephone. Therefore, the distribution of the sample cannot be tightly controlled. A personal interview can be made more leisurely, however, and many of these vital

TABLE 2
AUDIENCE BREAKDOWNS BY TELEPHONE AND NONTTELEPHONE HOMES
(Pulse Ratings)

	Telephone Homes	Nontelephone Homes
GREAT MOMENTS IN MUSIC WABC, Wednesday 10:00-10:30 PM	7.2	3.5
CRESTA BLANCA CARNIVAL WABC, Wednesday 10:30-11:00 PM	7.8	3.1
KAY KYSER COLLEGE WEAF, Wednesday 10-11:00 PM	15.8	20.8
MY TRUE STORY WJZ, Mon.-Fri. 3:15-3:45 PM	1.9	3.4
LISA SERGIO WQXR, Mon.-Fri. 9:00-9:15	2.0	.6

classification data can be recorded by the interviewer from her own observation.

In telephone surveys the respondent is likely to be surprised and confused by the unusual nature of the telephone call. It takes some time and explanation

before he can regain composure and answer intelligently. Furthermore, there is always the unknown factor about the status of the respondent in the family. His authenticity and acceptability as a respondent are seldom established. These are additional forces which raise doubts about the rapport established in a telephone survey but which can adequately be handled in a face-to-face situation.

The personal interview roster method of audience measurement ensures a balanced sample from quarter hour to quarter hour. Since each roster covers all the quarter hours within the time period being studied, and each respondent's listening for the entire roster period is recorded in one interview, a coherent and continuous flow of listening can be obtained, and comparisons between quarter hours in this roster period are valid because the sample is identical for each one. This is not possible in the telephone coincidental method of interviewing, which is primarily a record of instantaneous listening at the moment of the call. It is physically impossible to obtain absolutely balanced samples of respondents by telephone interviewing. Fluctuations in ratings of hour or half hour shows regularly occur.

Since the personal interview-roster method makes it possible to control the sample obtained, the equivalence of samples is assured, or at least their lack of equivalence is known. Thus the ratings of programs obtained by means of interviewing different but equivalent samples may be directly compared. On the other hand, random telephoning guarantees nothing about the equivalence of the various samples drawn. Comparisons of ratings of different programs are made with considerable injustice often done to programs because the particular samples drawn may favor one program more than another.

The personal interview-roster method obtains a record of "psychological" rather than mere physical listening to programs. Many listeners habitually tune out a given program when the concluding commercial begins, and turn to the station that carries the next program they want to hear. For a period of perhaps five minutes, while the station announcement and terminal commercials are being made, the respondent is psychologically "deaf," and cannot be considered a true listener until the program proper that he tuned in comes on the air. This deaf interval would be recorded and given equal value with the previous and following listening by the Neilson Audimeter, which is a mechanical device for measuring physical radio tuning. It would be wrongly recorded as *the* actual listening by a telephone coincidental interviewer who called after the original program was turned off and before the new one began. On Station Y, Program B is closing. The new tune-in is not psychologically listening to this program. He is not aware of its contents or its commercials and does not report it in the roster interviewing technique. Meter records this as if it is equal listening with that from 8:00-8:25 and from 8:30-9:00. Tele-

phone coincidental made at any moment in this period makes respondent aware of "listening" he is really not doing. In both instances incorrect reports are obtained.

DIAGRAMMATIC PRESENTATION OF PSYCHOLOGICAL "DRAFNES"
IN RADIO TUNING

8:00	8:30		
Station X Listeners to Program A			
	8:25 at 8:25, tune in Station Y to wait for Program C		
	Program B	Station Y Program C	
	8:25	8:30	

A personal interview would not record such physical tuning because the respondent would remember that he heard program A and then turned to Program C. Only by the aided roster recall method of interviewing does psychological listening assume its true value.

Interviewing by the roster method obtains a record of continuous listening over a period of several hours. Studies of flow of listening from program to program can be made, and various degrees of listener interest in specific programs or stations can be studied. The meter, which on the surface should also permit such flow studies, fails because its minute-by-minute account records meaningless dial-twisting and unintentional and deaf listening as discussed above. Thus, flow of audience would be incorrectly traced.

Criticisms of the roster method of interviewing frequently stress the fact that it relies on the respondent's memory, which is likely to recall "name" shows and network stations at the expense of lesser known programs or independent stations. These criticisms are true of unaided recall interviewing and of incomplete rosters, but experiments have proved that the method adapted by the Pulse in its regular surveys does not allow these inaccuracies to occur.

The Pulse roster is a complete listing of all available stations and programs by quarter-hour periods, as shown at the top of page 176.

Each interviewing day is divided into parts or roster periods, and interviewing for each daytime roster period is done immediately after the period closes, and for the night roster, at 9:00 A.M. the following morning. First the interviewer obtains a record of coincidental listening at the time of the

9:00-9:15 A. M.

WOR	V. H. Lindlahr, health talk
WABC	News—J. King
WHN	Gloom Dodgers
WINS	B. Chester Orch.
WMCA	News, Alice Hughes
WNEW	Dick Brown, Songs
WEAF	Mirth & Madness
WQXR	Angeline Dougherty
WOV	The Cipduzzas
WJZ	The Breakfast Club
WLIB	This is How
WHOM	Italian Variety Hou
WAAT	News, Nan Lupo
WPAT	Musical Scrapbook
WNYC	Masterworks Hour

interview [1] and she then proceeds to determine at what times, if any, during the period she is studying her respondent listened to the radio. Instead of relying upon unaided memory, a method that would obtain results distorted in favor "name" shows—the interviewer leads her to reconstruct her activities during the period in question.

After the respondent has determined at what times she had the radio on, the interviewer shows her a complete roster of the programs for each quarter hour mentioned, and asks which programs the respondent or other members of the family listened to at that specific time.

In this manner, for each roster period, completed interviews with responsible adults are obtained for every quarter hour period over seven days of the week. Thus in Philadelphia a total of 8,400 roster interviews are made, each of which measures listening over a period of 16 quarter hours or more. To obtain the same number of telephone coincidental calls for each quarter hour, a total of 201,600 telephone coincidental interviews are required.

In its published monthly reports The Pulse gives the following informa-

PHILADELPHIA, DECEMBER, 1943
6:00-6:15 P.M.

		% Homes	% Listeners
KYW	News, Music, Sports	2.1	10
WCAU	Robin Flynn Sports	2.5	12
WDAS	News, Day's Hits	.7	4
WFIL	Terry and the Pirates	4.4	21
WIBG	Danceland	2.4	12
WIP	War Extra	2.2	11
WJZ	News, J. Kennedy	1.9	9
WOR	Sydney Moseley	.5	2
WPEN	Bond Giveaway	2.6	13
Misc.		1.4	7
Sets-in-Use		20.5	100

tion: a "rating" or per cent of homes of the total number of interviews for each weekday quarter hour; a weekly rating based on a five-day total of interviews for each quarter hour from Monday through Friday; and a rating for each Saturday and Sunday quarter hour. A second figure indicating the per cent of listeners each program obtained is listed beside the rating.

The base in this instance was 5,000 radio homes. The first figure, or rating, was obtained by dividing the actual number of listeners to each station by 5,000, and the second figure, per cent of audience, was obtained by dividing the same number by 20.5, the sets-in-use figure for this particular quarter hour.

The application of the personal interview-roster method points to the ultimate in radio audience measurement. The collection of additional data facilitated by the personal interview makes possible numerous extensions of this technique:

(1) *Age, sex and family status of listeners.* A single question elicits this information whenever a home reports listening to the program in question. Consider these findings in connection with the Lone Ranger:

"THE LONE RANGER," WJZ, Mon.-Wed.-Fri., 7:30-8:00 P.M.

AGE AND SEX OF LISTENERS			
	%		%
Male	54	Under 20	38
Female	46	20-29	24
		30-44	40
		45 & over	8

It is apparent that the major part of this audience is composed of adults. It is likely that when children control the radio, the adults listen along with their children. However, there is no doubt that adults listen intentionally to this program too, and there are records of adult listening without the presence of children. It is unwise to presuppose the age and sex of an audience without any research data.

(2) *Socio-economic composition of audiences.* It is sometimes valuable to know what kind of an audience is reached by a program. The controlled sample obtained by the personal interview-roster method can be tabulated to reveal size of audience in the socioeconomic levels, and in telephone and non-telephone homes. The figures below are an example of this type of analysis:

"FRED WARING" PROGRAM, WEA, Mon.-Fri. 7:00-7:15 P.M.

SIZE OF AUDIENCE BY SOCIO-ECONOMIC LEVELS	
	% of Homes Listening to Fred Waring
A	11.7
B	6.5
C	3.6
D	3.3

Whether this is a good audience or not depends upon the intention of the advertiser. Certainly the mass of the consumers of this program's sponsored product is in the C and D group. Yet there may be other reasons for designing a program which may appeal to selected groups.

(3) *Other attributes of the radio audience.* In what kind of home do the listeners live? What about educational level and occupational status? Do they have an automobile or not? All those can be obtained in the personal interview. Here is an example of listening by automobile and nonautomobile families:

"VOICE OF FIRESTONE," WEAF, Monday 8:30-9:00 P.M.
SIZE OF AUDIENCE IN AUTOMOBILE AND NONAUTOMOBILE FAMILIES

	<i>% of Homes Listening to Voice of Firestone</i>
Automobile Homes	8.0
Nonautomobile Homes	4.0

(4) *Depth interviews with bona fide listeners.* It is often desirable to know how listeners react to a program. Do they like certain features and dislike others? What are their reasons for their listening? It is possible to compile a list of homes which listened to a particular program and to revisit them and conduct a depth interview. This has been done in numerous instances. For example, in March, 1944, a musical program featuring a popular local band and a short morale-stimulating talk by local leaders was broadcast over several stations located in an area dominated by a particular industry and inhabited largely by people of a particular national extraction. Was the program a good one? The answer was sought not only in terms of ratings indicating audience size, but also in reaction to the music and talk features of the program. A roster survey was made. Whenever a listener to this program was found, this listener was questioned further to determine his attitude and acceptance of the music and talks featured. The responses indicated a satisfactory reception of the program. Here are a few typical comments referring to the musical feature:

Our Frankie loves the Polkas
My husband wants to go to the broadcasts—he's crazy over Polkas
It's good—I'm waiting to hear Jolly Jack and the gang when they
will be at the theatre
I enjoy Polish music more than anything else

and here are several referring to the talk feature:

They are very interesting—to hear soldiers who are back from the front.
It's all about the industry and what they are doing. I think they are very
good and interesting.

Sometimes people I know talk. Once our Priest was on.
Whenever I listened they were interesting.
Very good yesterday when the soldier from the hospital talked.

(5) *Adjacent listening.* The listening behavior of a program's audience during the previous and following quarter hours or over an extended period is often revealing as to the particular appeal of a program. Does it gain a large new audience or "tune-in"? Does it "steal" listeners from other stations? Does it simply benefit by following and being placed in a chain of good programs on one station? For example, let us consider the Hit Parade audience:

ADJACENT LISTENING^[2] TO "HIT PARADE"
WABC, 9:00-9:45 P.M.

	<i>Audience comes from</i> 8:45-9:00 P.M. %		<i>Audience goes to</i> 9:45-10:00 P.M. %
Hobby Lobby	21	WABC Sat. Night Serenade	60
Truth or Consequences	23	WEAF National Barn Dance	2
Swap Night	8	WJZ News—J. McDonald	4
Others	7	Misc. Others	5
Not listening	41	Not listening	29
Total Hit Parade Audience	100	Total Hit Parade Audience	100

In this case, two significant events are observed. The program succeeds in creating a huge new audience which was not listening to the radio prior to the program. It also attracts a somewhat larger share of its audience from Truth or Consequences on WEAJ rather than from Hobby Lobby on the same station. The psychological reason for this may be the common bond in the Hit Parade and in Truth and Consequences. It is the contest element present in the two programs offering a challenge to the listener.

(6) *Sales effectiveness.* The best commercial measurement of radio listening will be not simply listeners, but listeners who buy. An example of this type of measurement is shown in the following study of the sales effectiveness of the Jack Benny program in relation to Grape Nuts Flakes. Two equivalent groups were interviewed with respect to the purchase of this product. The groups differed only in the fact that the first group listened to Jack Benny, while the second group did not.

	<i>Total</i>	<i>% Buy Grape Nuts Flakes</i>
Jack Benny Listeners	254	9.4
Nonlisteners	366	5.5
Difference		3.9 (additional purchasers in each group of 100 listeners).

The program rating is 41.0. This means that in the total market of all possible buyers of Grape Nuts Flakes, 41 per cent listened to the program. Therefore the net increase in buyers due to the radio program is computed by multiplying 3.9 (the additional purchasers in every 100 listeners) by 41.0 (the number of listeners in every 100 possible buyers in the market). This figure is 1.6, the number of additional buyers secured by radio in every 100 possible buyers in the market. [3]

This analysis is effective and becomes meaningful only with repeated measurements. The single instance does not afford a sure basis for interpreting causal relationships. However, if repeated investigations indicate that among listeners there is a rising use of the product, whereas among the nonlisteners there is not, then causal interpretations may be made. In this case, a repeat made some time later indicated that among nonlisteners it had declined slightly to 5.3 per cent.

In the end, the measurement of the radio audience should be made in the unit of listeners who buy. A frequent complaint against low ratings is the point that the advertiser is satisfied because it is selling his product, and the low rating must therefore be wrong. This is not correct reasoning. It is more likely that the low rated program in this case is not carrying "excess baggage." A large proportion of the listeners must be responding by purchasing the product.

(7) *Duplication in listening.* A sponsor has a program on Station X at 5:30 P.M. He is interested in a second program on Station Y at 12:30 P.M. The question naturally arises whether the second program will reach a new and additional audience. The type of program is different and the time of day is different. It is expected that it does reach a different audience. Such a question can be answered when the two programs are in the same roster. In this particular case there was practically no overlapping in audience of the total audience in the two programs, only 3 per cent listened to both of them.

The above applications of the personal interview roster method, measuring the radio audience, point to the expanding developments in radio research. It is not stopping with "nose counting" alone.

REFERENCES

1. For a comparison of telephone coincidental interview ratings with personal interview-roster coincidental ratings, see: Roslow, Sydney, *Measuring the Radio Audience by the Personal Interview Roster Method. Journal of Applied Psychology*, 27, No. 6, (December, 1943).
2. Fishman, Sidney, and Roslow, Sydney. *The Study of Adjacent Listening in Radio Research* 1942-3, 397 ff. Ed. Lazarsfeld and Stanton, New York, 1944.
3. This is a method developed and reported by Rensis Likert in the *Journal of Applied Psychology*, 20, No. 2 (April, 1936), 175-182.

Chapter XIV

THE RADIO PANEL

SAMUEL G. BARTON
Industrial Surveys Company

THE marketing, advertising, and radio broadcasting industries have for several years been equipped with several indexes of the relative popularity of local and national radio programs.

Advantages of speed, accuracy, flexibility, and economy had firmly established, by 1939, audience measurements based on telephone contacts. Direct questions determined to which program the respondent was listening at the moment his home was called.

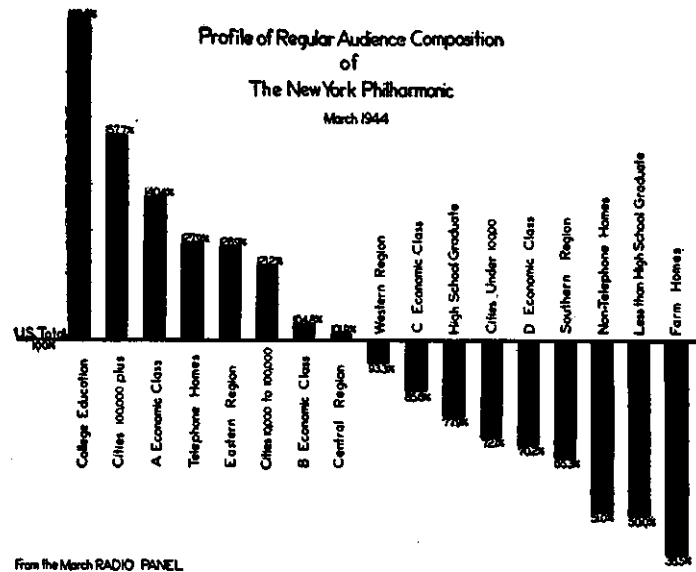
Limitations of the method were apparent. Studies were made (primarily in local areas) using field interviewers in an effort to include nontelephone homes as well as telephone families. Methods for measuring other audience dimensions also were sought. These included estimates of the net cumulative total radio homes or people listening at any time during the program, and a measure of the distribution of families or people according to the number of minutes of listening during any given period.

In addition to technical difficulties, the practical limitation of cost was a major deterrent on collection of data through personal interview, despite the gains to be derived from a broader universe and the possibility of studying audience characteristics and movement (see Charts I and II).

The cost of direct interview also appeared to impose limitations on the extent of geographical area that could be economically surveyed resulting in a tendency of the industry to employ the field interview technique primarily in larger urban communities and in the more densely populated counties lying within a short radius of the city limits.

It was well known that the listening areas of any two radio stations are not coextensive. This means that any arbitrary selection of counties close to city boundaries would (by varying degrees) be unrepresentative of the exact listening area or potential total audience for any station.

CHART I



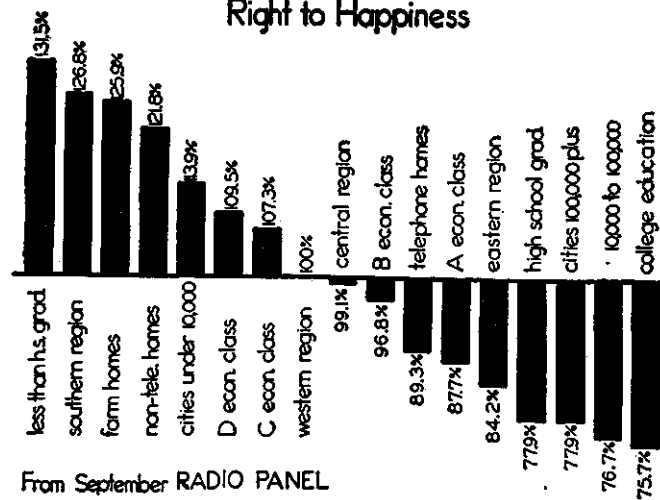
To construct the above visualization of the characteristics of the New York Philharmonic audience, we have taken the proportion of families listening to one or more broadcasts of the Philharmonic during March in the United States as 100. Proportion listening in each subdivision of the United States is then expressed as a ratio of the United States average.

While the above limitations were perhaps of minor importance, the desirability of measuring certain other radio audience dimensions resulted in the common need for a third method of data collection and a third technique of measurement.

The radio field needed measures of audience turnover *between multiple broadcasts* of a program series, together with analysis of the nature of flow of audience from quarter hour to quarter hour between competing programs. A strong need existed for more facts about audience composition: sex, age, educational level, income, and other individual listener characteristics. An extension of these requirements included the need for data about the value of listeners as consumers of any given type of commodity. It was felt that in many cases the nature of the program was poorly correlated with entertainment preferences of the type of person whom the advertiser hoped to reach.

During 1939 the research staff of Industrial Surveys Company began a series of experiments applying scientific treatment to the generally misused and rejected procedure of collecting data by mail. The treatment involved the following basic factors:

CHART II
 Profile of Audience Composition
 of
 Right to Happiness



From September RADIO PANEL

The above visualization of audience composition of the radio program, Right to Happiness, was constructed by taking the proportion listening to any one broadcast in the United States as 100. The proportion listening in various subgroups of the United States is then expressed as a ratio of the United States average.

1. Predetermination of the sample based on random or quota selection within population strata.
2. Permanent identification of the sample selected by means of a control or serial number indicating stratum, block and cell number and group sequence.
3. Compensation of the respondent, either *in advance* of work completed or upon return of the desired data depending on the nature of the study.
4. Repetition of contact with variation of incentive or compensation to those members of the sample who had not returned a previous mailing.
5. Double sampling or subsampling of those members of the sample not responding after completion of the series of repetitive mailings. The subsampling of nonrespondents being conducted by direct investigator contact.

A battery of experiments resulted in the development of the Controlled Mail Method of data collection. It was first applied nationally on the Columbia Broadcasting System county-by-county station audience survey conducted during the spring of 1940. In this survey, which included every county of the

United States, a national return of 83.4 per cent was achieved. A minimum return of 50 per cent was developed in every county and a return of 70 per cent or more obtained for each state.

Technically, this study also demonstrated that on surveys where geographical distribution of the sample is essential, cost per case is lower for controlled mail than for collection by field investigators. A degree of co-operation was developed which exceeds the actual degree of representation obtained in the usual door-to-door survey. This is due to the high proportion of individuals and of entire families not available to the interviewer in the usual one call survey made during normal working hours.

Employing the controlled mail system of data collection, Industrial Surveys staff then began experiments in the application of *panel technique* using both the recall and the diary method of obtaining program listening data.

Panel technique, while long used in other research fields, had not been applied with any success in consumer research. This was due primarily to the high cost of continuous collection of data from the same sample. The development of the controlled mail system made the panel technique a valid form of data collection which would permit a repetitive and continuous contact with the same sample of participating people. Panel technique *requires daily or weekly collection of data* for a high level of accuracy. This, however, results in a relatively high field cost. In turn the volume of detail which must be processed to obtain summary or average measurements for any period of listening time is far greater than that found in the nonpanel procedures.

In 1940 the research staff conducted the Des Plaines Study. All radio homes in the city of Des Plaines, Ill., were used in the concurrent comparison of all commercially used radio audience measurement techniques.

The coincidental (momentary recall) technique with data collected by telephone interviews among homes so equipped, and by personal contact among nontelephone homes, was taken as the standard for comparison. The following techniques and methods were compared to the standard:

DES PLAINES STUDY—1940		
	<i>Technique</i>	<i>Method of Data Collection</i>
Standard:	Coincidental	Telephone and direct personal interview
	1. Recall—aided	Direct personal interview
	2. Recall—aided	Controlled mail
	3. Recall—unaided	Direct personal interview
	4. Recall—unaided	Telephone interviews
	5. Recall—unaided	Controlled mail
	6. Diary	Investigator*
	7. Diary	Controlled mail

*By either method of collection, the actual record of listening was prepared by the participating family. At each pick up it was checked by the interviewer with some member of the family.

The purpose was to obtain a measure of the relative accuracy of the various techniques and methods, and to develop an accurate basis for comparing the costs of measuring various spans of listening time. Though techniques were found to show levels of total listening volume which varied beyond sampling error, agreement as to the distribution of listening by station ran high. Rank correlation of station by time periods averaged over .98 with the correlations tending to be lower between recall techniques and standard coincidental.

Recall techniques had a strong tendency to *underreport* the listening to programs of relatively low actual popularity and to overstate, to a smaller degree, programs of high actual popularity. No significant differences were found in distribution or level of audiences between methods of data collection within the same technique.

Through use of comparable cost tables it was concluded that each method held commercial advantages dependent upon the specification of the research problem. For studies involving short spans of listening time, the diary was found to be the most expensive per quarter hour and least expensive for surveys where an extended time span is to be covered.

Several valuable technical conclusions were also drawn which formed the basis for employing additional sample controls in subsequent controlled mail-panel-diary studies.

In October, 1942, Industrial Surveys staff began an extension of the experiment on panel technique on a national basis, directing a roster type aided recall survey to its established National Consumer Panel using controlled mail. In the same year, in co-operation with Columbia Broadcasting System, a further test of the diary-controlled mail method was made at Washington, D.C. Concurrently, a telephone coincidental survey was conducted each day from 9 A.M. to 10:30 P.M. during the week of the test.

The results clearly reflected the benefits in accuracy derived from the advance sample control of factors found to be highly correlated with listening volume. These included:

- A. Number of radios in the home.
- B. Number, sex, and age distribution of members of the participating families.

These factors, among others, are controlled in the sample selected for tabulation of the actual listening reports. The findings of the various tests were also of material aid in our operation of the 1942 and 1944 CBS county-by-county radio station audience surveys.

In 1943 Industrial Surveys Company began the continuing panel radio surveys published monthly as the *Radio Panel*. This series covers network programs. Panel studies were also begun in local market areas and have become known as "Listener Diary Surveys."

During 1944 the *Radio Panel* was converted from a bimonthly measurement to a weekly measurement and the Listener Diary Surveys have begun repetitive operation in several smaller markets as well as in most principal cities.

The types of audience measurement being drawn from the panels include:

1. Total audience listening to a specific broadcast (net cumulative—one broadcast).
2. Average audience to a series of broadcasts (Monday—Friday average) (four week average) (six-month average), etc.
3. Net weekly audience (proportion listening to any one broadcast of a multiple-day program).
4. Net monthly audience (proportion listening to any one broadcast during a four-week period).
5. Net 13-week audience.
6. Audience Turnover—weekly, monthly, annual, etc.
7. Distribution of families by number of broadcasts heard during any number of weeks or months.
8. Duplication (between two or more programs of the same sponsor or between competing product's programs).
9. Audience flow (source and dispersion of the audience to any program).
10. Audience composition by age and sex.
11. Audience composition by geographic location, size of community, economic class, educational level, and occupation of family head.
12. Radio-magazine-newspaper correlations.
13. Program-commodity purchase correlations.
14. Program-product correlations.

During the period of years when the controlled mail panels were being developed, the A. C. Nielsen Company and others were experimenting with mechanical means for recording the operations of radio sets in homes. Their developments in recording devices have been directed toward many of the needs for additional audience dimensions outlined above and for the additional measurement of minute by minute flow, turnover of audience within a program period, and the distribution of families by number of *minutes of listening*. The limitation as to dimensions, measurable by the recorders, evidently is solely that of omission of measurement of individual person listening and the present inability to function in homes unequipped with 110-volt electricity. Broadly, the objectives of the Panels and the recording devices are the same.

In this connection it is worthy of note that despite some criticism which has been leveled against the recorder technique on the basis of showing the

radio in use when not being heard, companies receiving reports of both the recorders and the Panel have found close agreement between the two. They report that the two series when tabulated for the *same sample geographical area* agree both as to total level of listening and as to distribution of audience by programs for most hours of the day and most days of the week within the limits of standard sampling error.

LIMITATIONS OF THE PANEL

While the controlled mail panels can be based on samples controlled by income and purchasing power, the self reporting technique limits the representation of the system to *literate* homes. This limitation has little effect on the volume of reported listening to most programs. It does result in a tendency to overestimate somewhat the audience to the few programs which have their greatest appeal in upper social classes, and to understate somewhat those with strong appeal in the *illiterate* brackets. These limitations are obviously minor deterrents in view of the system's ability to represent literate families with battery-powered radios as well as radio families in electrified homes.

DIFFERENCES IN THE USE OF BATTERY RADIOS

The panel measurements have indicated marked differences in the use of radio between rural homes with battery radios and rural electrified homes. From nature of the listening pattern it has been concluded that owners of battery radios are conserving their batteries during the war time shortages. This results in consistent peculiarities in the *amount* of listening and in selection of news programs with radios being turned on and off shortly before and after the news broadcast.

During normal times, other researchers have found major differences between the listening to battery and all-electric radios effected by the two factors of tendency to conserve the batteries and secondly the difference in quality of reception of the two types of receivers.

With the increasing attention being given to programming by both large and small broadcasters and the possibility of accurately studying the fitness of the program for the market and job it is assigned to do, it is likely that the controlled mail-panel system will be widely employed during subsequent years for qualitative analysis and control.

Chapter XV

THE SURVEY APPLIED TO EDITORIAL PROBLEMS

H. C. LUDEKE

The Curtis Publishing Company

THERE is an increasing use of the sampling survey to solve problems which hitherto have been considered the province of the journalist or literary man. This increased use is particularly evident in the field of magazine publishing.

A distinction should be made between editorial research engaged in by a writer to gather economic, historical, or social data for inclusion in a particular article or story; research designed to verify the accuracy of statements made by writers who have submitted manuscripts to a periodical for publication; and research designed to guide editors and publishers in the construction of periodicals which will yield greater satisfaction to the prospective reader. We are here concerned only with the last-named type of research. This is essentially "product" research. Its purpose is to ascertain the likes and dislikes of the market (prospective readers) and to generalize this information so that it becomes a working tool for the editor. Wherever we use the term "editorial research" it refers to this product marketing operation.

Because editorial research is ordinarily concerned with the attitudes of large numbers of people, it usually must rely for its data on the sampling survey rather than upon a complete census. This is especially true of periodicals of national circulation running into the millions, but it is also true of sectional papers and those catering to selected professions, trades, or social groups. As a rule it is important to study the nonreader as well as the reader, and this implies a sampling of whole populations.

The methods of editorial research are basically the same as those of market research, or, more broadly speaking, those of social research. They differ only in surface aspects as the nature of the problem calls for this variation or that, in order to make the tool usable.

The application of the scientific method to editorial problems has been retarded by a feeling on the part of some that editorial problems were so unique, so much a matter of creative spirit that they could not be analyzed and studied coldly with a "nuts and bolts" approach. One editor is reputed to have said, "There are at least a million people who feel as I do. If I make this magazine to suit myself, it will please this million just as well." Modern editors no longer take this view. They realize that there are many aspects of editing which can be planned in advance and many decisions made on the basis of ascertainable facts instead of by guesswork and intuition.

The job of editing still, however, is an art and probably always will be. Even after the editor has the facts he has to know what to do with them. His is still the task of assembling into a single composite unit all the editorial items that go to make up the magazine. Many considerations go into the selection of the separate items and the form of their presentation, such as titling, illustration, position, and the area they are to occupy. Additional considerations affect the *number and kind of items which are to be assembled and the manner in which the whole is to be "packaged."*

The alert editor knows that he needs all the help he can get on these matters. He doesn't shun research, doesn't feel he is editing "by slide rule" if he uses it. After all, every editor thinks he knows certain things to be true about his trade. Often this knowledge is merely folklore of the business and at best based on limited observation. Research merely substitutes for this loose sort of "knowledge" verifiable facts based on carefully observed and recorded data from an adequate number of cases. Such facts no more inhibit the individual expression of the editor than notated scales and theory hamper the work of a composer.

Many problems in the editorial field still defy solution by scientific means, but only because the proper tools for working on them have not yet been invented. Social research, to use the broad term, is still in its infancy. As new techniques and methodology are developed (and they are being developed constantly), and as we gain familiarity with our data, the areas in which the scientific survey cannot operate will diminish. Meanwhile there are countless editorial problems which can be solved by the sampling survey.

Let's take an actual example of a survey made *some time ago* at the request of the editor of *Country Gentleman*. We will take the problem as it was originally presented and show what steps were taken to provide the answer. The conditions which prompted the survey, of course, no longer hold, and the survey data have no current value. Nevertheless, we can see, in outline, the nature of an editorial problem as it is presented to the research man and the way in which such a problem can be handled by the survey method.

The editor had been in charge of the publication for about one year. During

this time he had made several minor changes: an increase in the amount of space devoted to illustrations, the addition of certain feature items or columns which appeared regularly from month to month, and a change in the cover illustrations from paintings to photographs.

Such gradual changes are the rule rather than the exception in editorial operations, and for very good reasons. Any publication which measures its readers in the hundreds of thousands or even millions holds those readers largely because they accept the publication as it is. Radical changes in the publication stand an even chance of being unwelcome to the current reader. Furthermore, such changes are immediately visible to the present reader, but they are not apparent to the nonreader whom the editor might like to win, until the changes have been brought to his attention by advertising, by sampling, through conversation, or merely by chance. In the ordinary course of events this is likely to produce a time lag which, from the short-term view, may actually result in a circulation loss.

In this instance the editorial changes had been made without benefit of research in advance of their adoption. Although they might have been pre-tested, the editor preferred to give these changes an actual trial run so that the reader could judge from actual experience over a period of time whether they were desirable or not.

The editor's prime concern therefore was whether or not these changes had met with favor or disfavor among current readers. At the same time he was anxious to have any additional information which would enable him to strengthen the publication's hold on this group.

The first step was to list the specific questions which the survey was to answer. Each member of the editorial staff contributed. From this list a preliminary plan of action was drawn up. After field pre-testing, the final issues of the study were agreed upon:

1. Which type of covers is preferred, photographs or art work?
2. With the growth of pictorial journalism is the magazine lagging behind the demands of readers in the amount of space devoted to illustrations?
3. Most of the editorial space is devoted to factual articles rather than fiction. Should the current relationship be changed?
4. Length of articles has been cut down. Do readers approve? Do they want a return to the longer article or do they want them still shorter?
5. Serials or continued fiction stories are expensive compared to short stories per unit of space they occupy in the magazine. What is the relative popularity of serials and short stories?
6. Are women readers of the magazine satisfied with *Country Gentlewoman*, the special section devoted to homemaking interests? Is it

properly balanced with respect to the space devoted to the various aspects of homemaking? Is the material practical?

7. One of the regular farm features dealing with crops, livestock, poultry, etc., has been given a pictorial treatment in various issues of the previous year. Do the readers approve?
8. Several new departments have been added to the magazine. Are they being read? Are they useful to the reader?
9. What is the over-all evaluation of the magazine compared to that of its competitors?

A twofold approach was decided upon: questionnaire and observation. The specific questions of the editor were incorporated into an interviewing schedule (see Fig. A and B). So far as possible the reader was to be presented with alternate choices specifically stated. The number of opportunities for a "free response" was purposely limited. The pre-test had demonstrated that the average person has difficulty in expressing his likes and dislikes in editorial material. To some extent this is perhaps a matter of having given the matter little thought. To a greater extent it is a matter of vocabulary. People are aware of the net impressions the magazine makes upon them, but they cannot define the elements which produce it. Over and above this is a feeling of awe for things editorial. There is a tendency to take for granted things seen in print. This causes reluctance in suggesting how the magazine might be changed. A good deal of the inarticulateness was overcome by bluntly putting questions: "Which will you have? This, this, or this?"

In addition to the questionnaire we also planned to secure a record of observation and reading of every item of editorial content in three consecutive issues of the magazine. This record was to be gotten from readers of the magazine. It would be a record of what they had *done* as contrasted with how they *felt* about the questions on the schedule.

It was necessary to work out the details of how the information was to be gathered. With readers located in every state of the Union the use of a mailed questionnaire would have been very convenient. However, the danger of a majority of the replies coming from the more ardent readers and from the better educated was too great to permit serious consideration of this method, even if there had not been other considerations. It was decided the only feasible method was the personal interview.

One thousand interviews had to be so distributed that they would fairly represent the entire readership of over two million copies of the magazine distributed monthly from coast to coast. Since over nine out of ten copies of the publication were sold on a subscription basis, the problem could be narrowed down to sampling subscriber homes. The distribution of sub-

WOMEN

A STUDY OF THE COUNTRY GENTLEMAN

	X	CODE
1. Here are two sets of covers from the Country Gentleman. One set represents the 1940 series and the other the 1941 series. Which of the two types of covers do you like better?		
1940 _____		25-1
1941 _____		2
Both Equal _____		3
Don't Know _____		4
Why do you choose this type? (Do not offer these answers but check as many as are given.)		
Like paintings better than photographs _____		26-1
Like photographs better than paintings _____		2
Like the subjects better _____		3
Look more natural to me _____		4
Look more like a farm magazine _____		5
Picture can be cut out _____		6
Other: (Please specify) _____		

Don't Know _____		7
2. How do you feel about the amount of pictures used to illustrate Country Gentleman articles? Would you like:		
MORE PICTURES AND <u>LESS</u> PRINTED TEXT? _____		27-1
LESS PICTURES AND <u>MORE</u> PRINTED TEXT? _____		2
NO CHANGE? _____		3
Undecided _____		4
No Opinion _____		5
3. Thinking of the amount of fiction and articles in Country Gentleman as it is today, which would you personally like better:		
MORE ARTICLES AND LESS FICTION? _____		28-1
LESS ARTICLES AND MORE FICTION? _____		2
KEEP IT AS IT IS NOW? _____		3
Undecided _____		4
No Opinion _____		5
4. In general which do you like better in monthly magazines:		
SHORT STORIES? _____		29-1
CONTINUED STORIES? _____		2
Undecided _____		3
No Opinion _____		4
5. As a rule, which do you like better:		
FICTION WITH AN HISTORICAL SETTING? _____		30-1
FICTION WITH A PRESENT-DAY SETTING? _____		2
Undecided _____		3
No Opinion _____		4

		WOMEN	
		X	CODE
6a. Do you read the movie column "Facts About Current and Coming Motion Pictures?"	Regularly		31-1
	Sometimes		2
	Never		3
b. How do you feel about it?	Like		32-1
	Neutral		2
	Dislike		3
	Undecided		4
	No Opinion		5
c. Have you ever used it in deciding whether or not you (or some member of your family) would see a certain movie?	Yes		6
	No		7
	Don't Know		8
7. What do you think of the way Country Gentlewoman handles these subjects?			

	Practical			Amount of it?			
	Yes	No	No Opinion	Too Much	About Right	Too Little	No Opinion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Recipes			33				
Other Food and Cooking			34				
Household Appliances and Equipment			35				
Dress Patterns			36				
Ready Made Dress Fashions			37				
Housekeeping Suggestions			38				
Furnishing and Decoration			39				
Beauty Helps			40				
Childhood Problems (Under school age)			41				
Gardening			42				

		WOMEN	
		X	CODE
8. Which of these things would you have a keen interest in reading about?			
	Weaving _____		43-1
	Antiques _____		2
	Crocheting _____		3
	Women in the News _____		4
	Knitting _____		5
	Book (reviews) _____		6
	Making quilts _____		7
	Music _____		8
	Making party decorations and favors _____		9
	Sports _____		0
	Making toys _____		X
9. Do you happen to take an active part in any community activities?			
	Yes _____		44-1
	No _____		2
If "Yes," what kind?			
	Club _____		45-1
	Church _____		2
	School _____		3
	Grange _____		4
	Other: (Please specify) _____		
10. Do you happen to read Farm Journal-Farmer's Wife?			
	Yes _____		45-1
	No _____		2
If "Yes:" a. If you could have but one, which would you choose?			
	Farm Journal _____		46-1
	Country Gentleman _____		2
	Either One _____		3
	Undecided _____		4
b. What in particular appeals to you in Farm Journal?			
	_____		47
c. What in particular appeals to you in Country Gentleman.			
	_____		48

scriber homes was compiled by rural cultural regions so that the relative importance of each, as a source of circulation, was available.

These rural cultural regions are those described by A. R. Mangus in the report *Rural Regions of the United States*, based on a study conducted by the Division of Research of the Works Project Administration, published in 1940. The advantage in using these regions, instead of the conventional census regions, lay in the fact that they are constructed on the basis of type of farming, standard of living, land value, tenancy, and race—all of which can be assumed to have a bearing on reading habits and tastes. Census regions on the other hand are drawn merely on political boundaries.

Cost considerations prohibited use of every cultural region. By confining the sample to those containing 3 per cent or more of the subscribers, we could, however, work with a group of fifteen regions which together accounted for 80 per cent of all subscribers. The counties which statistically most closely matched the regions in which they were located were chosen as the interviewing points.

Instructions for conducting the interviews were prepared in written form. Women interviewers were recruited and trained for their specific task by supervisors who were sent to the areas where the interviews were to be made.

Interviewers represented themselves as employees of the Editors' Research Council, a fictitious name used to avoid biasing those interviewed in favor of *Country Gentleman*. They were furnished with galley sheets of C.G. subscribers for the county assigned to them. Each was assigned a quota of interviews by towns and farm areas of the county and asked to interview the same number of men and women over 17 years of age. Otherwise, selection of names from the subscriber lists was to be "random."

The following quotation from the instructions will serve to explain the method employed to insure an unbiased selection of homes to be interviewed:

In filling your quota of farm homes, first determine by talking with local farm authorities (County Agent, etc.) which is the *most typical* rural route out of town. By "typical" we mean those with farms near the average in worth and engaged in raising the crop typical of the vicinity. Then sort out the *farm* names on the list which are located on the rural route. Call on these homes as they come—regardless of their appearance or type. This will be made easier for you if, before you start, you have had the help of some local person in arranging the names in the order in which you should come to them on the road.

In following this general procedure you will have to exercise some judgment. For example, if you find that calling at every home on the list will fill your entire quota before you have gone more than a few miles from town you should skip every second or third name, etc., so that you get a good spread over the whole route. If, on the other hand, every name on this typical route is still insufficient to fill your quota, start out on the next most typical route, etc. Keep in mind that the reason behind any routing plan is to include in our work all kinds of farms in their proper proportion—big ones as well as little ones, good ones and poor ones, remote ones as well as those near town.

The same principle holds for selecting names to fill your non-farm or "in town" quota. Find out locally which is the *average* residential street (or streets) in town, *not the best nor*

9-10-11-41 EDITORS' RESEARCH COUNCIL

WOULD LIKE TO READ		ELEMENT NUMBER	SEEING				READING			
YES	NO		DID NOT SEE	SAW	CERTAIN		READ NONE	READ PART	READ ALL	DON'T KNOW
					REASONABLY SURE	UNCERTAIN				
1	2		3	4	5	6	7	8	9	0
25		1								
		2								
		3								
		4								
		5								
30		6								
		7								
		8								
		9								
		10								
35		11								
		12								
		13								
		14								
		15								
40		16								
		17								
		18								
		19								
		20								
45		21								
		22								
		23								
		24								
50		25								
		26								
		27								
		28								
YES	NO	ELEMENT NUMBER	DID NOT SEE	SAW	REASONABLY SURE	UNCERTAIN	READ NONE	READ PART	READ ALL	DON'T KNOW

WOULD LIKE TO READ		ELEMENT NUMBER	SEEING				READING			
YES	NO		DID NOT SEE	SAW	CERTAIN		READ NONE	READ PART	READ ALL	DON'T KNOW
					REASONABLY SURE	UNCERTAIN				
1	2		3	4	5	6	7	8	9	0
53		29								
		30								
55		31								
		32								
		33								
		34								
		35								
60		36								
		37								
		38								
		39								
		40								
		41								
65		42								
		43								
		44								
		45								
70		46								
		47								
		48								
		49								
		50								
75		51								
		52								
		53								
		54								
		55								
80		56								
YES	NO	ELEMENT NUMBER	DID NOT SEE	SAW	REASONABLY SURE	UNCERTAIN	READ NONE	READ PART	READ ALL	DON'T KNOW

4-	6-	8-	9-				
5-	7-		10-				

1. Was checked copy acquired primarily for this respondent?	X	Code			
Yes		11-1			
No		2			
2. About how long has respondent been a reader of checked publication?					
Less than two years		3			
Two years or more		4			
3. About how often does respondent read a copy of checked publication? (Approximately)					
Every issue or every other issue		5			
Less frequently		6			
4. Does respondent intend to read more in <u>this</u> issue?					
Yes		7			
No		8			
5. Respondent's Sex?					
Male		12-1			
Female		2			
6. Respondent's Color or Race?					
White		3			
Other		4			
7. Respondent's Age?					
10 to 14		13-1			
15 to 19		2			
20 to 24		3			
25 to 29		4			
30 to 34		5			
35 to 39		6			
40 to 44		7			
45 to 49		8			
50 to 54		9			
55 to 59		0			
60 to 64		X			
65 & Over		Y			
8. Relationship to Family Head?					
Family Head		14-1			
Wife		2			
Daughter		3			
Son		4			
Other - Male		5			
Other - Female		6			
9. Highest Grade of School Completed by Respondent?					
		15-			
10. Data on Home:					
On Farm		16-1			
Not on Farm		2			
Owned by Family		3			
Rented by Family		4			
11. Family Economic Status?	X	Code			
Farm					
Non-Farm					
F. \$40 & Over per wk.		17-1			
M. \$30 to \$39 per wk.		2			
L. \$15 to \$29 per wk.		3			
VL. Under \$15 per wk.		4			
12. a. Chief Source of Family Livelihood: Farming					
Farming		5			
Other		6			
b. If "Farming" describe kind					
		18-			
		19-			
c. If "Other" describe occupation and kind of business.					
		20-			
		21-			
13. Respondent's Employment:					
Housework		22-1			
Student		2			
Retired		3			
Farmer		4			
Other (specify occupation and kind of business)					
		23-			
		24-			
Resp's Name	_____				
Rural Route or St. & Number	_____				
P.O. or Town	_____				State _____
Publ. Checked	_____				Issue _____
Date of Interview	_____				
Interviewer's Signature	_____				

198 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

yet the poorest. Pick your names at random from those streets. Do not, however, pick or even attempt to pick the *homes* on the list which you or anyone else thinks is that of a representative family. This usually results in selection of a family well *above average*.

Each interviewer carried with her copies of competing publications as well as of *Country Gentleman* in order to give the impression that more than one magazine was included in the survey. They also carried a portfolio containing the following C.G. covers:

<i>1940 Issues</i>	<i>1941 Issues</i>
January	May
May	August
July	September

Her first step was to explain her association with the Editors' Research Council and then to show the copies of the magazines, asking "Do you ever see these magazines?" If the answer was "No," she asked to see someone else in the house.

If the answer was "Yes," she asked if the person had by chance looked into any one of these particular issues. If the person had looked into only the competing magazines, the interviewer excused herself on the ground that her quota of those interviews was complete.

If the person had looked into that issue of *Country Gentleman*, the interview was started by spreading the portfolio of C.G. covers before the respondent for cover preference. The other questions then followed.

It may be of interest to look at some of the results of this study, and later to see what action was taken because of them. A complete analysis of all the questions on the schedule or of the results of the page-by-page checking of the three copies of the magazine is beyond the scope of this chapter. We can, however, examine briefly certain of the more significant points.

The 1941 photographic covers proved considerably more popular than the 1940 painted covers. While several reasons were advanced for this preference, the outstanding and most logical reason was that the photographic covers "looked more natural."

TABLE 1
CHOICE BETWEEN 1940 AND 1941 C. G. COVERS

	Total 100%	BY SEX		BY AGE		
		Males 100%	Females 100%	Under 30 100%	30-44 100%	45 & Over 100%
Total Answering:						
<i>Choices</i>						
1941	61	70	54	65	62	60
1940	21	13	28	26	22	19
Both Equal	15	14	16	9	13	17
Don't Know	3	3	2	—	3	4

MEM

A STUDY OF THE COUNTRY GENTLEMAN

1. Here are two sets of covers from the Country Gentleman. One set represents the 1940 series and the other the 1941 series. Which of the two types of covers do you like better?

	X	Code
1940		25-1
1941		2
Both Equal		3
Don't Know		4

Why do you choose this type? (Do not offer these answers, but X as many as are given.)

Like paintings better than photographs		26-1
Like photographs better than paintings		2
Like the subjects better		3
Look more natural to me		4
Look more like a farm magazine		5
Picture can be cut out		6
Other: (Please specify)		

Don't Know _____ Y

2. How do you feel about the amount of pictures used to illustrate Country Gentleman articles? Would you like:

MORE PICTURES AND <u>LESS</u> PRINTED TEXT?		27-1
LESS PICTURES AND <u>MORE</u> PRINTED TEXT?		2
NO CHANGE?		3
Undecided		4
No Opinion		5

3. This year Country Gentleman in its regular monthly section on Crops, Poultry, Livestock, etc. called "Along the Farm Front" has been running more pictures and less text than formerly. How do you feel about this change? Would like:

RETURN TO MORE TEXT AND FEWER PICTURES		28-1
STILL MORE PICTURES AND LESS TEXT		2
KEEP IT AS IT NOW IS		3
Undecided		4
No Opinion		5

200 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

		MEN	
		N	Code
4.	Thinking of the amount of fiction and articles in Country Gentleman as it is today, which would you personally like better:		
	MORE ARTICLES AND LESS FICTION? _____		29-1
	LESS ARTICLES AND MORE FICTION? _____		2
	KEEP IT AS IT IS NOW? _____		3
	Undecided _____		4
	No Opinion _____		5
5.	This year Country Gentleman has been publishing shorter articles than formerly, but there have been more of them in each issue. How do you feel about this?		
	WOULD RATHER HAVE ARTICLES MORE COMPLETE (EVEN IF IT MEANS FEWER OF THEM) _____		30-1
	WOULD RATHER HAVE ARTICLES STILL SHORTER AND MORE OF THEM IN EACH ISSUE _____		2
	KEEP IT AS IT IS NOW _____		3
	Undecided _____		4
	No Opinion _____		5
6.	In general which do you like better in monthly magazines:		
	SHORT STORIES? _____		31-1
	CONTINUED STORIES? _____		2
	Undecided _____		3
	No Opinion _____		4
7.	As a rule, which do you like better:		
	FICTION WITH A HISTORICAL SETTING? _____		32-1
	FICTION WITH A PRESENT-DAY SETTING? _____		2
	Undecided _____		3
	No Opinion _____		4
8. a.	Do you read the movie column "Facts about Current and Coming Motion Pictures"?		
	Regularly _____		33-1
	Sometimes _____		2
	Never _____		3
b.	How do you feel about it?		
	Like _____		34-1
	Neutral _____		2
	Dislike _____		3
	Undecided _____		4
	No Opinion _____		5
c.	Have you ever used it in deciding whether or not you (or some member of your family) would see a certain movie?		
	Yes _____		6
	No _____		7
	Don't Know _____		8

MSW

	X	Code
9. a. Do you read the new department "The Farm Engineer?"		
Regularly _____		35-1
Sometimes _____		2
Never _____		3
b. How do you feel about it?		
Like _____		36-1
Neutral _____		2
Dislike _____		3
Undecided _____		4
No Opinion _____		5
c. Do you find the information useful?		
Frequently _____		6
Occasionally _____		7
Never _____		8
10. a. Do you read the new department called "Handy Work?"		
Regularly _____		37-1
Sometimes _____		2
Never _____		3
b. How do you feel about it?		
Like _____		38-1
Neutral _____		2
Dislike _____		3
Undecided _____		4
No Opinion _____		5
c. Do you get ideas from it you can use?		
Frequently _____		6
Occasionally _____		7
Never _____		8
11. Do you happen to read The Farm Journal?		
Yes _____		39-1
No _____		2
If "yes:" a. If you could have but one, which would you choose?		
Farm Journal _____		40-1
Country Gentleman _____		2
Either One _____		3
Undecided _____		4
b. What in particular appeals to you in Farm Journal?		
_____		41-

c. What in particular appeals to you in Country Gentleman?		
_____		42-

Men were somewhat more partial to the photographic covers. Older persons were less inclined to choose between the two types, but those who did choose showed the same strong preference for the photographic covers shown by the younger people.

TABLE 1
REASONS FOR PREFERRING 1941 COVERS

	Men	Women
Total Answering	100%	100%
<i>Reasons for Preference</i>		
Look more natural to me	61	53
Look more like a farm magazine	55	49
Like photographs better than paintings	36	40
Like the subjects better	31	28
Picture can be cut out	4	10
Other reasons	3	1

While 1941 covers ostensibly appealed because of their "naturalness" the subjects portrayed could well have played a prominent part in the choices. For example "Looks more like a farm magazine," one of the most common reasons for preferring them, is quite possibly an expression of approval of the *subject* illustrated. If so, then these reasons in combination with the clear-cut references to subject matter would make subject matter the outstanding attribute of the 1941 covers.¹

Most regular readers of a publication are probably satisfied with it. The people interviewed in this study were for the most part regular readers of *Country Gentleman*. Therefore, in answer to the question:

"How do you feel about the amount of pictures used to illustrate *Country Gentleman* articles? Would you like more pictures and less printed text? Less pictures and more printed text? No Change?"

it is not surprising that 61 per cent answered "Keep it as it is" (Table 3).

TABLE 3
RATIO OF PICTURES TO TEXT IN ARTICLES

	BY SEX			BY AGE		
	Total- 100%	Males 100%	Females 100%	Under 30 100%	30-44 100%	45 & Over 100%
Total Answering:						
<i>Attitude</i>						
"More Pictures Less Text"	25	23	27	32	32	21
"Less Pictures More Text"	11	15	8	5	9	13
"No Change"	61	59	62	62	56	63
"Undecided"	1	1	2	—	2	1
"No Opinion"	2	2	1	1	1	2

¹ The average person gave more than one reason for his preference, which accounts for the fact that the figures in Table 2 total to more than 100 per cent.

Furthermore, this is a *safe* answer. It is difficult to say in advance how you would feel about a change, especially when there is no opportunity actually to see the nature of the changes proposed. Nevertheless, we can probably assume that among the majority there was *no active demand* for the changes suggested.

On the other hand, the *most acceptable* change is clearly an increase in pictures at the expense of printed text. This proposal was considerably more acceptable to women than to men and considerably more acceptable to younger persons than to older.

Only persons who had been readers of *Country Gentleman* for at least two years were queried about the Fiction-Article Ratio. Their attitude is indicated by the answers to the question:

"Think of the amount of fiction and articles in *Country Gentleman* as it is today, which would you personally like better—more articles and less fiction or less articles and more fiction?"

TABLE 4
FICTION-ARTICLE RATIO

	BY SEX			BY AGE		
	Total	Males	Females	Under 30	30-44	45 & Over
Total Answering:	100%	100%	100%	100%	100%	100%
<i>Attitude</i>						
"More Articles Less Fiction"	23	19	18	22	24	23
"Less Articles More Fiction"	19	16	22	30	19	18
"No Change"	54	51	56	46	54	55
"Undecided"	2	2	2	-	1	2
"No Opinion"	2	2	2	2	2	2

A small majority, as Table 4 shows, suggested the ratio be kept as it is now. That this majority should be somewhat less than on previous questions can possibly be explained on the basis that interest patterns on fiction and on articles vary sharply with age and sex. Previous research indicates this to be so. Consequently, any publication offering both fiction and articles to an audience varied as to age and sex will have some difficulty in arriving at a Fiction-Article Ratio that suits everyone.

Among the men approving a change, more would prefer that change to involve an *increase in articles* than an increase in fiction. Women interested in changing the ratio would, on the other hand, prefer an *increase in fiction* although the women were much more evenly divided on this point. In fact it would be safer to conclude that as many women advocated an increase in articles as advocated an increase in fiction. This is particularly interesting because it is generally felt that women are more "fiction minded" than men.

In the matter of age it appears that the older the reader the less the desire

for a change in the present Fiction-Article Ratio. The group most actively interested in a change were those under 30 years of age, and the change they suggested was toward fewer articles and *more fiction*. The older group, with less interest in change, however, voted for *increased emphasis on articles*. This agrees very neatly with what we believe to be youth's greater interest in fiction and the older person's greater interest in articles.

In Table 5 we examine the question of the Fiction-Article Ratio from the standpoint of residence and from the standpoint of primary versus secondary readership. The reader living on a farm indicated a slightly greater interest in a change in the ratio, and the change favored seems to be to *increase articles* at the expense of fiction. The reverse seems to be true of readers not living on farms although in neither instance are there clear mandates for one course or the other.

TABLE 5
FICTION-ARTICLE RATIO

	By Farm Residence		By Primary-Secondary Readers	
	Farm	Non Farm	Primary	Secondary
Total Answering:	100%	100%	100%	100%
<i>Attitude</i>				
"More Articles Less Fiction"	26	15	26	18
"Less Articles More Fiction"	18	22	18	22
"No Change"	52	58	53	55
"Undecided"	2	2	1	2
"No Opinion"	2	3	2	3

It seems worth mentioning that primary readers,² while giving a majority vote to *no change*, also preferred the *increase in articles* to a somewhat greater degree than they did *increase in fiction*.

Only those men who had been reading *Country Gentleman* two years or more were presented with this question:

"This year *Country Gentleman* has been publishing shorter articles than formerly, but there have been more of them in each issue. How do you feel about this? Would you rather have articles more complete, even if it means fewer of them? Would you rather have articles still shorter and more of them in each issue or keep it as it is now?"

The "no change" vote fell just short of winning a majority, but on the whole, those willing to see a change seemed to prefer shortening the articles still more (and having more of them). This, however, seems to reflect the

²The terms "primary readers" and "secondary readers" will require explanation. As we have used the term, a primary reader is the person in the home whose interest in the *Country Gentleman* is primarily responsible for its being there. Other readers are considered "secondary." The primary reader is, therefore, a very important person to the editor.

TABLE 6
SHORTER ARTICLES
(Only men who have read *Country Gentleman* 2 years or more)

	Total 100%	BY AGE			BY FARM RESIDENCE	
		Under 30 100%	30-44 100%	45 & Over 100%	Farm 100%	NonFarm 100%
Total Answering:						
<i>Attitude</i>						
"Prefer More Complete"	19	28	17	18	21	14
"Prefer Still Shorter"	32	31	36	32	33	29
"No Change"	46	35	43	48	43	54
"Undecided"	1	6	1	1	2	--
"No Opinion"	2	—	3	1	1	3

preponderance of older persons in the group interviewed. The younger group showed more inclination for a change, but they were almost deadlocked on which direction the change should take.

Those living on farms show more inclination for a change than those not living on farms and seem to prefer further shortening of the present length of articles. Town people seem better satisfied, but evidently those willing to see a change prefer to see shorter articles.

What does an editor do with information of this sort? One thing is fairly certain. He doesn't rush out to the printing department and remake the next issue then and there. He will more than likely proceed with considerable caution, all the more so if his publication has been successful. It is risky to quarrel with success, but the progressive editor will constantly make improvements in line with the desires of his market and in line with the objectives he has set for himself.

Changes in a magazine may also be retarded by previous long term commitments or by altered circumstances in the market such as the advent of war. In this case there seemed to be a preference among *Country Gentleman* readers for covers illustrated with photographs. This idea was not crystal clear, but it was a fair assumption. However, several months before the survey was made, contracts had been signed with twelve American artists, to paint a series of harvest scenes, each typical of a different region of the country. These twelve paintings were to serve as cover illustrations for the twelve consecutive issues of 1942.

These paintings were run as scheduled in spite of the survey. They probably would have been run even if the survey had been made earlier; it is conceived to be the duty of an editor to lead as well as to follow. The proposed paintings apparently offered an unparalleled opportunity to bring to rural America some of the finest examples of native contemporary art. The editors believed that *Country Gentleman* readers should have this opportunity. They exercised

their right to lead, and their judgment seems to be vindicated by the fact that there was subsequently an unprecedented demand for reprints of these covers. Still, three-quarters of all covers run in the following year were illustrated with photographs.

Another point brought out in the survey was that if any change was made in the picture-text space ratio of articles the most acceptable one would be to increase the proportion of space devoted to illustrations. This, it will be remembered, was not the will of the majority, but of the minority who were willing to see some change. Hence it would have been folly to have made sweeping changes in this direction. Instead, there has been a slight reduction in the proportion of article space devoted to text.

TABLE 7
PROPORTION OF "FRONT OF THE BOOK" ARTICLE SPACE DEVOTED TO TEXT*
(Pre-Survey Year 1941 Considered as 100)

	<i>Index</i>
1941 (Pre-Survey Year)	100
1942	102
1943	99
1944 (1st 6 mos.)	91

*"Front of the Book" refers to the section between the lead article and the first advertising page.

Table 7 shows the tendency of the editors to reduce gradually the proportion of article space devoted to text and conversely to devote a larger proportion to illustrations for "Front of the Book" articles. These articles include general interest articles on broad national problems as well as articles on more specific problems of farming. It is a section of high interest to men readers, and it will be recalled that men proved less interested than women in an increase in the space devoted to illustrations.

TABLE 8
PROPORTION OF COUNTRY GENTLEWOMAN ARTICLE SPACE DEVOTED TO TEXT*
(Pre-Survey Year 1941 Considered as 100)

	<i>Index</i>
1941 (Pre-Survey Year)	100
1942	91
1943	92
1944 (1st 6 mos.)	87

*Country Gentlewoman section devoted nearly entirely to homemaking advice.

In Table 8 the tendency to reduce the proportion of space devoted to text is also apparent. Here the reductions are somewhat more pronounced than

they were in the "Front of the Book" section, which is in line with the greater acceptability of such a change among women readers.

A sizable number of readers had shown an interest in an increase in the total space devoted to articles, at the expense of a reduction in fiction.

TABLE 9
PROPORTION OF EDITORIAL SPACE DEVOTED TO FICTION
(Pre-Survey Year 1941 Considered as 100)

	<i>Index</i>
1941 (Pre-Survey Year)	100
1942	81
1943	83
1944 (1st 6 mos.)	86

Since the entire space of a magazine can be thought of as devoted to either articles or fiction, any increase in the proportion of one is at the expense of the other. Table 9 indicates that between the pre-survey year of 1941 and the first six months of 1944 there occurred a 14 per cent reduction in the proportion of space devoted to fiction.

One of the most difficult editorial tasks is to reduce the number of words in a manuscript written by a skilled writer so that the given subject is covered adequately, clearly, and without too much loss of the writer's style. In the survey we are discussing, readers expressed a desire for shorter articles.

TABLE 10
AVERAGE LENGTH OF COUNTRY GENTLEMAN ARTICLES
(Pre-Survey Year 1941 Considered as 100)

	<i>Index</i>
1941 (Pre-Survey Year)	100
1942	106
1943	110
1944 (1st 6 mos.)	108

Table 10 proves that this desire was not fulfilled in the period immediately following the survey. There were, however, extenuating circumstances. This post-survey period found the country engaged in war, and agriculture became one of the critical industries. The editors felt a keen sense of responsibility for the job of keeping their segment of the farm population well informed. Farmers not only had to be convinced of the need of increasing the production of foodstuffs, but they had to be instructed in methods of raising special new crops which they had never before raised. They needed information on methods of increasing total production in the face of a dwindling supply of farm labor. The dangers and methods of avoiding inflation of land values had

208 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

to be pointed out. As a consequence, material which in normal times might have been more condensed could not, in the editor's opinion, be abridged without the risk of losing comprehensiveness or understandability.

These examples of survey findings and the action taken on them will serve to show how research has functioned on one editorial problem. We have seen how pertinent information can be supplied to an editor. We have also seen that the application of these findings is the editor's function. He may accept them or reject them according to the situation in which he finds himself. The essential thing is that whatever he does, he does in the full light of the facts. The survey method materially reduces the editor's uncertainty. If he adopts a course which is contrary to the wishes of his readers, he knows that he is doing just that. And we have seen that an editor does not edit "by slide rule" when he uses research. He can still *lead* his readers, but he knows where he is going, and he knows how many are following.

Chapter XVI

HOW MANY READERS?

CORNELIUS DU BOIS

Life Magazine

CHAPTER XV in this volume deals with "readership" measurement in terms of thoroughness of reading editorial matter, and Chapter IX concerns readership of advertisements in the magazine. The present chapter is concerned with measures of readership in terms of the number and types of readers themselves as ascertained in surveys comparing two or more magazines.

PURPOSES

The common aim of all such studies is to delineate the magazine's market, or some aspects of it, sometimes for the publisher's guidance, but usually for the edification of advertisers.

Many of the basic facts about magazines can be known without any market research at all. From the Audit Bureau of Circulations, a user of magazine advertising space can obtain reliable facts showing how many copies of each issue were sold, by what means they were sold, how much the reader paid. He can also obtain the geographical distribution of the sales of a typical issue by states and by city size; and the larger publishers at least are accustomed to providing the circulation figures in still more detail, showing the sales in individual towns and counties. For trade and technical publications, audited analysis of reader occupations is also available.

But advertisers are an inquisitive lot, and the questions they ask seldom stop with the audited circulation figures. They want to know the answers to questions such as these: What is the duplication of reading between Magazines A and B? What opinions do grocers have for Magazines M and N? How many readers per copy are there for Magazines X and Y? They can tell from the sales figures which magazines are relatively strong in big cities compared to rural

210 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

areas, but they need to know how the magazines compare with each other in reading among young women or young men, or among owners of automobiles, or among the youngsters who will be tomorrow's consumers.

To answer questions like these has required all the ingenuity that market research technicians could muster.

PROBLEMS

Magazines are edited for states of mind, and hence their audiences are among the hardest things to measure accurately.

Sampling. When you are trying to measure the market for a brand of tooth paste you have to design an accurate sample because the users of tooth paste vary in different parts of the country. But the variations in tooth paste users are as nothing to variations in magazine reading. If you talk to college graduates, you may find that four in ten have read an issue of a certain magazine, but if you talk to those who never went beyond grade school you may find less than one in ten.

This sampling problem becomes still more acute in the case of magazines because of the nature of the research assignment. In many market studies it does not matter much whether you find 15 or 20 per cent of the people doing the things you are trying to measure. It often does not matter whether you find 33 per cent of the upper income group and 18 per cent of the lower income group saying your product is worth buying, or whether you find that the figures are 38 per cent and 23 per cent. Whatever the actual figures, you will have demonstrated a difference, provided clues to the exploitation of a market. But in magazine measurements, as in the public opinion polls, the aim is often to provide a numerical measure. You cannot be satisfied in learning that Magazine A has more readers than Magazine B, because your assignment is to find out just *how many* readers there are with the smallest possible margin of error. And since magazine reading varies so greatly with different elements of the population, it becomes overwhelmingly important to make sure that the sample is representative for the varying elements of the population. Some of the more important variables which must be controlled are geographical (section of the country, city size, distance from market center, etc.). Others are personal characteristics (sex, age, race). And, most importantly, it is necessary to control cultural and economic factors (income or economic level, and education).

Prestige. The prestige factor poses a particular problem in magazine measurement. It is true, of course, that a certain number of people will claim to use tooth paste because they think it is the proper thing to say. This factor is much more likely to inflate claims to magazine readership. Among quality

magazines, more people than actually read the *Atlantic Monthly* or *Harper's* claim to do so. And in the mass field, this is the case with *Life*, the *Saturday Evening Post*, or *Good Housekeeping*, and a number of others. Conversely, a kind of negative prestige operates with some important magazines. People read them, but do not wish to admit it.

Confusion. The element of confusion creates another measurement problem. The confusion that exists in people's minds between two issues of the same magazine will invariably be greater than that existing between two brands of tooth paste or of breakfast food. While this confusion is not nearly as great as the confusion between two advertisements of the same advertiser, it is a factor which has to be taken into account, because a significant number of people will honestly think that they have seen the January 25 issue of a given magazine when in fact it was some other issue that they saw.

METHODS

The sampling factor, the confusion factor, and the prestige factor must all be kept in mind while evaluating the prevalent methods of measuring magazine readers. Only with these in mind can the meaning of the figures that result from each method be properly assessed.

Although the measurement methods are almost infinitely varied, they fall into three broad categories which can be labeled *objective* measurements, *attitude* measurements, *memory* measurements.

Objective measurements. Almost without exception the methods which can be grouped as objective measurements represent various attempts to circumvent the confusion and prestige factors.

The favorite method of this type is the magazine inventory. The usual procedure is to get into the house on some pretext or another and make a record of what magazines are actually found in the home. Physical presence of the magazine is the only acceptable evidence—claims made by respondents are not considered. (Some practitioners, however, modify the technique to include respondents' statements that certain magazines have been in the home, but have since passed along to others.)

This technique does not measure reading by the individual, and it misses those people who were exposed to the magazines outside their own home. But it does enable the researcher to describe the kinds of homes in which different magazines are found and in which they are presumably read.

Under certain circumstances the inventory technique contains a major weakness. While ostensibly completely objective, it can and often does include a sampling bias inherent in the method, a bias caused by the prestige factor. This was demonstrated several years ago when a magazine inventory was

attempted in connection with *Life's* Continuing Study of Magazine Audiences (see page 215). At that time, in addition to the usual audience measurement, it was decided to include a magazine inventory. About two thousand of the interviews included this feature, all of which had subsequently to be discarded. They could not be used even as part of the regular audience sample. The audience figures obtained from these particular interviews went skyrocketing compared to other audience figures obtained on the same magazines at about the same time with the same kind of sample allocation.

Further investigation in the next wave of interviews showed clearly that there is a big difference in both quality and quantity of reading between the people who will invite you into the house to look at magazines and those who will allow themselves to be interviewed on the doorstep or elsewhere. Those who read very few magazines or who read only those with negative prestige can find plenty of excuses for not permitting a magazine inventory.

This by no means invalidates the inventory completely. If the magazine inventory is conducted as part of a larger study, so that magazine prestige cannot affect the cooperation or non-cooperation of respondents, then the inventory technique can be counted as valid and useful.

In trying to remove the prestige bias from the inventory technique and to circumvent the stigma factor which affects certain magazines, some practitioners have given the technique an added fillip. In this variation an interviewer comes to the door offering to pay cash for back numbers of certain magazines. This method has worked with pretty fair success, unearthing copies of magazines about which people do not like to boast. But, on the other hand, it can understate the penetration of magazines which have a real prestige, with which people do not care to part for a pittance.

Some magazines, particularly those with a large subscription circulation and little or no newsstand sale, have studied their readers more or less objectively through subscriber panels. Information about the subscribers is obtained by interview, by direct mail, or sometimes by credit reports. In the last case the subscriber does not even know that he or she is a member of the panel. This technique has been useful to individual publishers in learning facts about the age, sex, income, education, occupation, etc., of readers. But most such panels are conducted individually for one magazine, thus making it difficult, if not impossible, to compare magazines by this method.

In a further attempt to be objective, several people have been experimenting lately with mechanical or chemical gadgets to measure the reading of magazines or the traffic through them. These are similar in intent to the machines which measure the use of the radio. So far none of these experiments have borne public fruit, but there is reason to believe that they will do so in the not too distant future.

An account of an experiment along similar lines was published by *Time* in 1934. In this experiment a professional fingerprint expert was retained. Calling at the homes of subscribers, the expert and his staff gave each subscriber a fresh copy of the magazine in exchange for one which had been in the home for about a week. Back in the laboratory, the magazine was subjected to an elaborate chemical treatment to develop the latent prints on the paper. Each copy was then analyzed primarily for the volume of traffic through it, as indicated by the number of prints on each page. As a by-product, the expert discovered the prints of 3.26 identifiably different people inside the average copy.

Objective enough to be classified as such is the consumer diary technique, in which a carefully selected panel of people is paid premiums to keep records of all magazines entering the home. Confined to magazines, this technique is subject to the same weakness as the inventory. As part of a larger study, its usefulness increases. The Industrial Surveys Company uses this technique in its national consumer panel, in which the primary reason for paying people is to keep a record of their grocery and drug purchases.

Attitude measures. The second major category of magazine measurement is the attitude study. Under this heading it should be recognized that even the question of "what magazines do you read regularly" is in essence an attitude question. People's definition of what constitutes reading and what constitutes regularly may vary greatly between individuals. And the same individual, influenced partly by prestige, may even interpret the term differently for different magazines.

When a magazine enjoys such prestige that too many people claim to read it, it is in a fortunate position indeed. The question might arise whether a magazine is entitled to exploit this prestige by using questions to bring it out. The answer, of course, is "Yes"—provided the results of the questions are presented in prestige terms, and not in terms of absolute measurement of coverage.

In such general questions, whether they are done by personal interview or by mail, the magazines with more prestige will usually profit more from free answer questions than from check list questions, although in most of the paired tests of the two approaches, the rank order has been about the same.

Some of the attitude questions which have been frequently asked are:

What magazine do you prefer?

What are your favorite magazines?

What are your first and second choice magazines?

Of all magazines published in the U.S. today which do you consider the most important?

214 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

In which magazine are you most likely to read what's on every page including the advertisements?

What magazine do you get the most out of?

If you could have only one or two magazines which would they be?

Which magazine do you think is best read by your customers?

Which magazines would you like to have your company advertise in?

As you can see from the questions themselves, they are usually designed to bring out the strong points of one or more magazines, and they should almost always be taken to represent an opinion rather than a fact.

Very often in such attitude measures the researcher (or promoter) is interested in only one or two questions. Such measures therefore lend themselves to direct-mail rather than interview questioning. The samples for such studies are usually more or less homogeneous, such as attempts to get the opinions of Pontiac dealers, or customers of the J. L. Hudson Company, or graduates of Smith College, or doctors under 45 years of age, or business executives of certain kinds of companies. To reach such lists is not only feasible, but also more economical by direct mail.

There is always, of course, the question as to whether those who answered by mail are typical of the whole group. It is undoubtedly true that those who will answer a mail questionnaire about magazines are better readers than those who will not bother to reply. Tests have shown that the competitive rank order of magazines among the nonanswerers conforms to the rank order among those who do answer. Since the purpose of such studies is not to get a numerical measure, but to get competitive or comparative ranking, the direct mail technique is usually considered valid and appropriate.

Memory measures. The third category of magazine measures is based on memory. Although these methods are also subject to prestige and confusion factors, they have the virtue that they do not accept mere *yes* or *no* answers regarding magazine readership. They attempt to classify the person as *reader* or *nonreader* on the basis of memory of items within the magazine.

Least successful of these methods is the attempt to get people spontaneously to recall stories or articles in a certain issue. This method will bring out the impact of the very strong story, but it counts *too much* on people's memories to serve as a reliable measure of readership.

Others have tried using a descriptive check list of stories, or table of contents. These methods ignore the fact that magazines are visual and that memory of magazines is also largely visual.

Another device is to show a portfolio of covers asking people which they remember having seen. But this, on the one hand, may pick up people who have merely seen the cover on display and, on the other hand, may miss people who did read the issue, but did not happen to remember the cover.

Then there is the device of a scrapbook of articles clipped from a magazine, sometimes including the cover as one page in the scrapbook. These clippings may be pasted full size; again they may be reduced photographically. Both have been used with pretty fair success, but they have the disadvantage that memory does *not* come in detached pieces. Showing people articles removed from their context may destroy some of the clues to memory.

The most successful method in this field is the one which has been adopted by the Magazine Audience Group¹ for the *Life* Continuing Study of Magazine Audiences (C.S.M.A.) and for the *Good Housekeeping* Study of Women's Magazine Audiences. Here the entire magazine is shown to the respondent, thus providing all possible cues to memory. But the respondent is not asked, "Did you read this issue?" He is asked, "Will you look through this copy and tell me which articles, if any, you definitely remember having seen?"

While this method provides accurate memory cues, it does not in itself get rid of the prestige and confusion factor. Consequently, the studies supervised by the Magazine Audience Group employed a device known as *confusion control*. This applies the same interviewing techniques to issues of magazines which have not yet gone on sale, measures the error, and corrects accordingly.

These *confusion control* interviews can also be a help in evaluating methods of questioning. For example, when the Continuing Study of Magazine Audiences was first conceived, it was felt that showing the cover of the magazine would probably increase the amount of confusion, because so many people would have a chance to see the cover, even though they had not seen the inside of the magazine. Accordingly, some 2,000 interviews were made in which the covers were stripped off the magazines, and a large scale *confusion test* was also conducted on this method. All the figures were too high. While some 6 per cent of the people were confused about the magazine with the cover on, over 15 per cent were confused when the cover was taken off. The "raw" uncorrected figures were also high. Obviously, in removing the cover, one of the cues to accurate memory had been removed. Consequently, the coverless method was abandoned, and all subsequent interviews were made with the magazine intact.

The weakness of the Magazine Audience Group's method is that it can be applied to very few magazines in one interview. But there are three such studies, two of which have reached publication, and nearly a dozen leading magazines are being separately measured in groups of three or four.

Because it has gained such acceptance as the standard of magazine audience measurement, the Magazine Audience Group method is worth further study.

¹ Consisting of Dr. Theodore H. Brown, chairman; Archibald M. Crossley, Dr. Raymond Franzen, Dr. George Gallup, Dr. Paul F. Lazarsfeld, Elmo Roper, and Dr. Samuel S. Wilks.

Here is the group's own description from a recent C.S.M.A. report:

The interviewing method finally developed is a very simple one. The interviewer carries with him one copy each of *Collier's*, *Liberty*, *Life*, *Satevepost*. These copies have been on sale approximately three and a half weeks at the time of interviewing—a time lag which tends to iron out the difference in the issue dates of the magazines and gives the total audience a chance to accumulate.

After the usual preliminaries in which the respondent is convinced that the interviewer has nothing to sell and no axe to grind, the interviewer asks: "Have you read or looked through any issues of these magazines within the last two months?" This lead-in question gives the blow-hards a chance to blow. If they want to impress the interviewer, or if they otherwise bias the answers, they can do it on this first question. For that very reason, the answers to this question are of doubtful value and are not presented.

The second question, the one we are really interested in, is worded as follows: "We are interested in learning whether or not you have read this particular issue. Because several weeks have elapsed since this issue was put on sale, would you check through and if you have read or looked at it, point out to me those pages you definitely remember."

Notice that the interviewer is not asking for a quick "yes" or a quick "no." He is asking the respondent to look through the issue before giving an answer, and to give the answer in terms of the pages which he or she saw. The same procedure is followed with each of the four magazines in rotating order, regardless of the answers to the first question.

After the first report was issued, this method of questioning was further intensified, and interviewers were instructed not only to get respondents to look through the issue, but to get them to go through the entire issue, page by page. This refinement resulted in greater accuracy, as indicated by the smaller amount of confusion which resulted.

All the rest of the questionnaire is made up of census and economic data for the purpose of checking the sample and for cross-tabulation.

Some of the newer interviews were made with a modification of this technique in which respondents were asked to do exactly the same thing—to go through the entire issue—but to give their answers in terms of individual items (stories and cartoons) in the magazine rather than in terms of pages. Severe tests showed that the results of this technique could safely be combined with the old technique and could be used in the future in cases where quicker interviews are desirable or where additional information is to be obtained in the same interviewing time.

To find out how much inflation must be removed from the answers, separate "confusion control" interviews were conducted with the same techniques. Three of the magazines are four weeks old at the time of interviewing; the fourth magazine has not yet gone on sale and will not go on sale for at least another thirty-six hours. Date lines on all four copies used in these "control tests" are unobtrusively masked. Each of the four magazines has been thus tested to the point where stable figures resulted.

In the "confusion control" all conditions are identical with the main survey, all

conditions except one: the magazine being tested has not yet appeared; the public could not have seen it, and any people who say they saw it, or something in it, are obviously inaccurate about this one magazine.² But remember that the control test contains this one impossible condition. It measures confusion before publication day (36 hours before).³ Subsequently the magazine *does* appear. If the same confused people were interviewed *after* the magazine appears, they will still say they had seen it, and most of their answers will still be wrong, confused answers—most, but not all. Some will really have seen the magazine and their positive answers will be correct. This is allowed for in the method of applying the “confusion control” as follows:

1. *Before publication* 4 persons out of 100 say they have seen something in Magazine A.

4 “yes”	96 “no”
---------	---------
2. *After publication*, all other conditions equal, 16 out of 100 say they have seen something in Magazine A.

16 “yes”	84 “no”
----------	---------
3. We subtract the confused persons, leaving twelve reliable “yesses” in a total of 96 reliable answers, or 12.5%.

12 “yes”	84 “no”
----------	---------

In the first three reports, this method of applying the confusion control was used to correct every single figure in every breakdown. Beginning with the fourth report, the above method was used only twice for each magazine, to correct the figures for adults and for children. These corrected figures are then used to tell the *amount* of confusion which should be eliminated, the page-by-page information being used to eliminate, in the proper quantity, the people whose answers were based on the least evidence. Following the example given above, this new method of application would be used to eliminate the 3.5% whose answers were based on the least evidence, to arrive at for Magazine A the 12.5% of people whose claims to reading Magazine A were based on the most evidence.

By this device the C.S.M.A. is enabled to tabulate not only a corrected *quantity* of readers; but also to tabulate in terms of individuals. Many detailed cross-tabulations can be run which would be impractical if every single figure had to be corrected by formula. Since adopting this device, the C.S.M.A. has become *more and more analytical*, and the number of factors which can be related to audience has greatly multiplied.

² But not necessarily about the others. There is no evidence that people confused about Magazine A are more than normally confused about Magazines B, C, and D. This gives us really a *control of our control*. The people who were confusion-tested about *Liberty*, for example, gave answers about the other magazines that conform satisfactorily with the results of the mail survey.

³ It was necessary in these tests to watch release date carefully in each community. Wherever newsdealers or boys were unofficially or officially selling copies early, or postmen were making early deliveries of subscription copies, we had to cancel interviews, and as a double precaution, all control tests were made at least 36 hours before the sale of each magazine.

Along these lines, those in charge of the study have recently spent considerable time in studying the apparent determinants of reading, by cross-tabulating with each other each of the factors which are significantly related to size of audience. This process has served to clarify a great deal of thinking. There is, for example, an apparent relationship between age and the reading of *Life*. The younger the person, the more likely he is to be a member of the *Life* audience. When age is cross-tabulated with education, however, this relationship disappears. The higher audience percentages in the younger age groups can be accounted for entirely by the greater educational levels attained by the younger people compared with the older ones. A similar cross-break of education and economic level again shows education to be the dominant factor, but in this case there remains a perceptible influence of economic level itself.

A study which aims so ambitiously at precision must take special safeguards. During 1944 and 1945 the Magazine Audience Group was engaged on behalf of *Life's* C.S.M.A. in a two-pronged sampling experiment aimed both at the problem of selecting communities and at the problem of selecting respondents within the community. It is constantly necessary to validate and revalidate the sample, to examine interviewing techniques, to study possible causes of error in the interview situation. A few years ago a question was raised in a Committee meeting as to the possible effect of interview fatigue on the audience figures. Accordingly, a special test was run to determine whether a magazine which comes fourth in the order of questioning suffers because of the order. As a test it was inconclusive: there did seem to be a slight tendency for audience figures to be lower at the end of the interview than at the beginning, but the difference was statistically insignificant. In order to play safe anyway, a system was adopted of rotating the order in which each magazine was shown to the respondents, according to a system developed by Professor Wilks, a member of the group.

To make sure that the interview procedure creates no more exaggeration and confusion than can be eliminated by the confusion control, several parallel tests have been run. In one of these scrapbooks were used containing articles from the current issue of the magazine and also from an advance issue so that the respondent could be tested simultaneously for true recollection and confusion. Preparation of these scrapbooks requires time, and the time is extremely short during which advance material from weekly magazines is available for interview. This test, therefore, could not be run nationally, but it was tested locally against an exactly parallel sample in which the regular C.S.M.A. technique was used. The results were nearly identical.

Because the confusion control device is both costly and complicated in operation, there have been constant attempts to devise an interviewing pro-

cedure which could be used nationally and which could eliminate confusion and exaggeration without the necessity of a separate control. During the study's first six years such a method had not been found. But the search for it went on during 1945.

The first step was to re-examine the causes of confusion in relation to this objective. As already mentioned, one was the prestige factor. It is a familiar problem in field research on a prestige product that however much the interviewer stresses that a "no" answer is just as valuable as a "yes," the average person tends to feel that "yes" is better. Some respondents, either to satisfy their egos, or in a misguided spirit of helpfulness, will claim reading they haven't done. Some outlet was needed, other than reading, for the desire to make a good showing in the interview.

The other important factor in confusion is the fact that similar pictures, articles, etc., do appear in different places. But again, the average respondent is not thinking about this angle; he is looking at a particular issue, and if he sees something familiar, he probably assumes it was seen in this magazine. Some way was needed of pointing up the difference between seeing something before, and seeing it *in this issue*.

With these requirements in mind, experimental questions were designed, tested on regular issues to determine their feasibility in the field. From these tests, various questionnaire versions were developed and tried out in a new series of confusion control tests. Altogether, 9,253 interviews were made in 1945 on advance copies of the weeklies to develop the questionnaire which yielded practically no confusion.

Here is the procedure:

First, an introductory question: "Have you looked through any copies of these magazines in the last 6 months . . . either at home or somewhere else like a friend's house, office waiting room, etc.?"

The old interview started off with a somewhat similar question. However, in the new questionnaire this introductory question covers three magazines in addition to *Collier's*, *Life*, and the *Post*. Thus, the respondent attaches no undue importance to the weeklies as against the others. People who say they haven't seen the weeklies in the last 6 months are dropped, so far as audience-measurement is concerned.

The second step in the old interview was to show specific issues. But in the new interview, people are given another chance to let their egos expand before asking about reading. They are asked what kinds of editorial material they like and whether they can remember any recent stories or articles from the weeklies.

And third, before producing specific issues, the problem of genuine con-

fusion is attacked by pointing out to respondents that similar material does appear in different magazines and newspapers, and asking: "Have you ever noticed that yourself?"

This completes the stage setting for the introduction of the magazines. Specific copies are then shown, but in a different manner than previously: The old approach was: "We're interested in learning whether you have seen this issue. Because several weeks have elapsed since it went on sale, would you check through, and if you have seen it, point out to me those items you definitely remember."

The new approach is: "Here's an issue which you may not have seen yet. Would you leaf through it with me and give me your first impressions of which items look interesting."

The interviewer follows up by asking two questions on each item: (1) Does it look interesting? and (2) Is this the first time you have seen it? This pattern of questioning is repeated until the respondent either has identified five items, or has been shown five without identifying any. In the first case, questions on interest are dropped, and the rest of the issue shown to see whether additional items have been seen or read. In the second case, questions on recall of editorial material are dropped, and the interviewer continues to find out which items look interesting. If a respondent who did not recall any of the first five items does identify something later on, this will, of course, be recorded.

In this way, the leafing-through procedure is constructive both from the reader's and the non-reader's point of view. The incentive to exaggerate reading is removed because the non-reader can talk about what would interest him.

Finally, items on which people are genuinely confused are recorded that way. For example, an answer such as "It looks familiar" goes down as "May have seen—not sure." "I've seen it somewhere, but I'm not sure where," goes down as "Seen before—not sure in this issue."

And there is a check-up at the end of each issue shown: "Just to keep the record straight, are you sure you have (or haven't) seen this issue before—or aren't you certain about it?"

Thus genuine confusion is filtered out, both for specific items and for whole issues.

It should be clear that this procedure applies only to people who say they have seen the weeklies in the last 6 months, and only to as many magazines as they claim to have seen. People who say they haven't seen the weeklies in 6 months are asked a separate set of general magazine questions. As usual, the final part of the interview consists of factual information needed for sample validation and cross tabulation purposes.

The effectiveness of this technique may be judged by the fact that whereas

the confusion remaining in the old method was large enough to have an effect on audience figures, the Magazine Audience Group has decided that no correction for confusion is necessary with the new method.

The remaining amount of confusion is so small that it could be accounted for by the fact that magazines are sometimes made available, and therefore can be seen, before the official publication date. In addition, special parallel tests of the old and new interviewing techniques indicate that the new method, in filtering out all but the most positive answers, eliminates a few people who have actually seen the issues being studied—in fact, that the method reduces the audience by slightly more than the number of confused people who remain.

Chapter XVII

MEASUREMENT OF EMPLOYEES' ATTITUDES AND MORALE

ARTHUR KOLSTAD
Houser Associates

INTEREST in surveys of employee attitudes and morale has increased greatly during the past decade. Management is giving more recognition to the importance of the relationship between what employees think and the efficiency of commercial and industrial activities.

Top management seldom knows how employees feel. Under stress of modern production it is often difficult for top executives even to keep in close touch with the foremen or supervisors. Lines of communication from management down through the various levels of supervision to the rank-and-file employees may be well developed. Lines of communication in the opposite direction are usually subject to breaks, interruptions, and poor connections. As a tool to aid management in securing a better understanding of how employees feel and think about their jobs, company policies and practices, quality of supervision, working conditions, and so on, the survey of attitudes and morale has been developed.

A number of different approaches to the problem of measuring employee attitudes have been used. Among these are the personal interview, either guided or unguided; the essay type of questionnaire, where the employees are asked to express themselves in their own words; the "Yes-No" or "True-False" questionnaire; the "Agree-Disagree" questionnaire, either with the two answers or with five answers providing degrees of acceptability to the respondent; the "Attitude Scale" consisting of a series of simple statements from which the employee is to select those with which he agrees or which he believes to be true; and the multiple-choice type of item questionnaire. In a number of studies combinations of some of these approaches have been used.

Each method has advantages and disadvantages, certain strengths and

weaknesses. The particular procedure must be selected in terms of the specific purpose or objective to be attained.

Experience with all of the above-named methods suggests that for general use in measuring employee attitudes and morale, the printed questionnaire with objective multiple-choice items is best.

The same principles governing the construction of any good questionnaire should be followed in developing an employee attitude questionnaire. These have been covered sufficiently in psychological publications and do not need review here. But certain points should be stressed. Since some employees who are somewhat deficient in schooling will be found in the average industrial concern, more than average care must be taken in formulating questions.

Questions should contain only simple, commonly used words.

The method of registering an expression of opinion should be made as easy as possible. The objective form of question which the employee answers by simply checking his choice among a number of ready-made answers meets this requirement.

An opportunity should be afforded for the expression of varying shades of opinion. The multiple-choice item with responses arranged in steps or degrees will provide this.

Topics to be included in the questionnaire will depend on the underlying purpose of the survey, the type of organization in which the survey is being made, the particular problems which confront management. In general, a study of employee attitudes should provide a measure of over-all job morale. It should also give an employee opportunity to express his attitude and opinion on a large number of specific points concerning his job; working conditions, his relationship to his fellow employees and supervisors, his understanding of management's policies.

Since the basic ingredients of the employer-employee relationship are essentially the same in all companies, most surveys will cover certain fields, such as quality of supervision, training and instructions, promotion policy and practice, job placement, and satisfaction with the job. The phases to be covered in these fields may vary according to the companies' policies. Though some questions can be used repeatedly with great value, the questionnaire as a whole needs modification and revision for each new study. In other words, the questionnaire needs to be tailor made to fit the particular organization.

The method of conducting the survey can have a real effect on the results and needs to be carefully controlled. In a typical survey, employees are assembled in conveniently sized groups in a room equipped with tables and chairs. Each group receives the same explanation and instructions, with emphasis on the point that the questionnaires must not be signed, that no one in the company will have access to the answered questionnaires, that there is no means

of identifying a particular person's blank. All of the mechanics of distributing the questionnaire forms and the placing of the answered forms in the ballot box are such as to guarantee anonymity to the employee.

As stated above, to be of greatest value to management provision should be made for securing some measure of over-all job morale. A single index summarizing the individual morale level, or the average group level is useful in making comparisons.

What is employee morale? It must be defined before an instrument to measure it can be developed. The rather broad meaning given the term must be broken down into specific elements about which the employee can be questioned. If we accept the concept that morale is made up of "a lot of little things," the number of elements to be included becomes very large. The sensible point of view calls for the selection of those phases which will provide a usable definition of general satisfaction. It must then, of course, be remembered always that the term "employee morale" as used in the study refers to the summation of the elements included.

Experience over a number of years has confirmed the usefulness of a series of broadly phrased items which ask in a variety of ways "How do you like your job?" and "Is this a good company to work for?" The responses to the ten questions which constitute this "battery" are assigned score values which can be combined into a "morale score" for each individual.

The simplest way of presenting this definition of employee morale is to list the attitudes and beliefs which characterize the employee who makes the highest possible "score" on the series of questions. He . . .

1. Believes his company is one of the very best of all companies as one to work for.
2. Thinks the management of his company cares more about the welfare of its employees than does the management of any other company.
3. Knows of no other company where he would rather work if he could get an equally good job elsewhere.
4. Blames himself and not the company for any dissatisfaction he may have experienced on the job.
5. Thinks the company does a great deal more than one might expect to promote good working relationships between himself and the people with whom he works.
6. Feels that he is really regarded as a part of the organization.
7. Thinks the management of the company is always fair with people in jobs such as his.
8. Thinks the people just above him are always fair in their treatment of him.

9. Knows of no other company that treats its employees as well as his does.
10. Feels that he can be sure of his job as long as he does good work.

The employee with very low morale would be the one who holds the exact opposite attitudes.

On this battery of items, the responses are assigned weights on the basis of the intercorrelations of the items and the frequency distributions of the responses, resulting in a possible range of scores of 0 to 100.

Scores of 0 and scores of 100 are rare. Most employees have attitudes somewhere between these two extremes. Examination of a number of survey reports shows that in practically all organizations studied, the majority of employees have morale scores somewhat above 50. This may be due to two factors: (1) surveys of employee attitudes are seldom made in companies which do not place considerable emphasis on employee relations and (2) employees in general can be expected to develop a degree of loyalty towards their own company.

The single index of morale for the entire company by itself is difficult to evaluate, unless the same method of measurement has been used in a sufficiently large number of companies so as to have some basis for comparisons. In the absence of such "norms" the frequency distribution of individual scores serves as a good indication of the morale situation in the organization. The single morale index becomes of considerable value in comparisons within the organization.

Some method of segregating the questionnaires by departments, divisions, sections, or other work groups, preferably through a departmental identification question, permits the comparison of the average morale scores of such groups. Differences observed between groups within a particular company have been larger than the differences between companies. In any organization having fifteen or more separate departments, it is not unusual to find a difference of as much as 25 points between the highest and lowest departments.

When considering specific attitudes, also, the basically important comparison is that between departments, divisions, sections or other organizational units. Most useful comparisons are secured by segregation of employees into groups which are fairly homogeneous in terms of type of work done and whose members are all responsible to a single supervisor. A careful definition of such groups is assumed in the discussion of the specific values and uses of the survey results which follows.

One of the greatest values of employee attitude surveys is that they provide an objective measure of personnel administration at the supervisory level. Progress is difficult without some measure of results. Lack of such measurement has been a real weakness in personnel programs. Most department heads

and foremen receive regular reports on their production, payroll ratios, costs, errors, and similar physical and financial measures of their performance on the job. They know that their superiors also see these reports and that they are judged on the basis of such reports. The personnel part of their job is seldom checked.

Survey results give individual foremen and supervisors a measure of their own strengths and weaknesses. Results provide a forceful motivation for self-improvement and for raising the level of their own personnel practice. In addition the findings may awaken management to the real nature of the supervisory problem and the importance of making it possible for supervision to do an acceptable personnel job.

Table 1 summarizes the attitudes of employees in fourteen departments in a relatively small manufacturing plant. It shows the relatively weak and strong points on a number of phases of supervisory practice. The favorable responses have been abstracted from the multiple-choice questions and reworded into positive statements of favorable attitudes. The definition of "favorableness" has been set at a level within reach of a company with strong supervision.

The relative standing of any one department in the company can be presented to the department head without revealing the standing of any other specific department. Such presentations show the department heads (except in the best department) that a higher level of performance of that specific supervisory phase is possible under the particular company's working conditions.

Measurement of personnel administration at the supervisory level shows the supervisory staff that management is watching the personnel side of their jobs. It gives industrial relations executives a much needed leverage in improving policy administration.

At times a further check in the low-morale departments reveals that certain supervisors lack the temperament or training for their work. A good production record but poor personnel administration may indicate that a supervisor should be transferred from a line job to a staff job. "Ratings" given the bosses by their employees may be helpful in making decisions regarding supervisory promotions and upgrading.

If additional training is needed, data of this nature provide useful material for training courses and foremen's discussion groups. Concrete evidence of employees' reactions to their supervision makes a forceful and challenging introduction to the discussion of the duties and responsibilities of the supervisory staff.

Management cannot assume that all its policies are administered uniformly throughout the organization or that they are equally well understood by all

TABLE I
 PERCENTAGE OF EMPLOYEES WHO EXPRESSED FAVORABLE ATTITUDES ON VARIOUS PHASES OF SUPERVISION

<i>Attitude</i>	DEPARTMENT													
	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>	<i>K</i>	<i>L</i>	<i>M</i>	<i>N</i>
My boss has a very good idea as to whether I am doing good work or not.	80	73	42	56	33	66	64	54	63	48	76	74	88	67
The people above me understand almost all, or all, of the difficulties of my job.	20	64	19	42	22	58	57	50	37	56	53	74	64	52
There is a clear understanding between my bosses and me as to just what work I am supposed to do and where my duties begin and end.	30	100	36	42	33	56	27	58	37	44	47	58	72	45
I rarely, or never, get contradictory or conflicting orders from people above me.	60	82	42	57	67	77	82	71	52	64	82	95	92	88
Employees in my department are rarely, or never "bawled out" or criticized when they do not deserve it.	60	55	19	51	22	56	82	67	70	64	76	100	72	88
When employees are corrected or have their work criticized, this is usually, or always, done in a friendly and helpful way.	70	73	39	63	22	61	55	46	70	66	82	100	96	88
When I do some unusually good work, I usually, or always, get recognition or praise for it.	50	55	22	22	22	38	55	12	37	46	29	63	56	36
There is no favoritism shown in my department.	70	73	33	30	56	51	64	46	56	52	47	89	72	67
When I am given new or special instructions on the job, they are usually, or always, clear enough and definite enough so that I know exactly what I am expected to do.	80	82	42	69	56	79	73	62	78	68	76	79	92	82
When changes are made in the way I am to do my work, I am usually, or always, told why this is being done.	60	45	33	46	33	54	27	38	59	58	82	68	92	73
When I go to my boss for information or help on a difficult work problem, I get all the help I need.	40	64	36	53	22	53	45	46	44	46	88	74	80	67

228 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

employees, or even by the foremen under whose direct supervision the employees work. Management may try to give the employees every consideration and may believe that its policies are fair and generous. But in the final analysis a policy is only as good as the employees think it is. In one manufacturing company it is the stated policy that an employee doing unsatisfactory work is to be warned before being dropped. The finding that less than one-fourth of the employees in a certain department believe that this is done indicates considerable divergence from company practice in at least one section of the plant.

When a majority of the employees in a number of departments believe that they would not have a reasonably good chance of getting a fair hearing and a square deal if they had any cause for dissatisfaction, it is time to review the method of handling grievances.

Questions dealing with pay and promotion may reveal "sore spots." In some cases the unfavorable reactions resulted from misunderstanding, lack of correct information about the company's pay and promotion policies. Instances have been found where some department heads in ill-advised attempts to keep employees happy have held out hopes for advancement contrary to company promotion policy or contrary to the probabilities that there will be vacancies in the better jobs. Instances have been found where the company's promotion policy was so vague that employee dissatisfaction could have been predicted.

Minor dissatisfactions are seldom revealed in the ordinary course of contacts between employees and their supervisors. If they are, they are often passed off by the foreman as an unimportant "gripe," to which he cannot pay attention because of his deeper concern for the production record. Top management may be the last to hear grumblings. Many petty irritations and minor dissatisfactions could be remedied easily. If permitted to grow, they may result in slowdowns, waste, poor quality of work, poor work attitudes, and in time may develop into major grievances.

It may not be possible to foresee all of the possible sources of petty grievances when constructing a questionnaire. Provision can be made for remarks and comments. A blank page at the end of the questionnaire with the invitation to write "any comments, suggestions, complaints, questions, or compliments that you may care to offer" usually brings forth a flood of comments which reveal the nature of any minor dissatisfactions present.

The survey of employee attitudes paves the way for the introduction of a new program. In developing any educational plan, the wise procedure is to determine first how well prepared the student is for the planned presentation. Just as the school examination, or the audit of accounts, or the inventory of merchandise may have a twofold purpose, so also does the study of employee

attitudes serve (1) in giving a measure of what has been accomplished and (2) in providing the basis for future planning by defining the present situation. This pertains particularly to orientation of the employee and to the company's reports to employees about financial and production progress.

The Personnel Department can check on its placement program through the responses of employees to items asking them how well they believe their jobs fit them, how much of their training and experience is utilized in their work, whether people with less training and ability could do their work, etc. The Personnel Department may be interested in knowing what the employees (especially the newer ones) think of the way they were treated when being hired, and whether they received the needed information about company rules, regulations, and personnel policies.

The Public Relations Department should be vitally interested in the survey results. Many public relations men have come to realize that it is difficult to sell a company and its social-economic policies to the public unless the employees themselves are sold on these same policies. If the members of the "company family" are not happy and informed, their friends on the outside will be misinformed. Good community relations have their foundation in favorable attitudes of employees. A survey of employee attitudes may, therefore, be of considerable help to the Public Relations Department as a check on the past program or as a guide in planning a new program.

It is also possible to secure a measure of the attitudes of supervisors, foremen, section heads, and (in the larger organizations) the top executives as well. A different form of the questionnaire must be constructed for the supervisory staff, since questions to be asked are somewhat different from those in the rank-and-file. Supervisors must be asked about their duties, responsibilities, and relationships to each other and top management. In some companies a different form for each managerial level is desirable. Certain questions pertinent to all levels can be used in each of the forms to secure a direct comparison of the attitudes of rank-and-file employees and of the supervisory staff. Among these are the more general attitude items, such as those used in the "morale battery" described earlier in the chapter, and any question dealing with relationship of an employee to his "superior," wherein the employee, whatever his level, has direct contact with a level above. Illustrations of the latter type of items are the following, applicable to both rank-and-file and supervisory:

When difficult problems come up in your work, how free do you feel about going to your boss for information or help?

How easy is it for you to see your boss if you wish to have a talk with him?

230 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

When you do unusually good work, do you get recognition or praise for it?

How well do the people above you understand the difficulties of your job?

When you are given new or special instructions on the job, how often are they clear enough and definite enough so that you know exactly what you are expected to do?

Though in general the attitudes expressed by supervisors have been more favorable than those of rank-and-file employees, many exceptions have been noted. In nine out of thirty organizations recently compared, a larger proportion of the rank-and-file employees than of the supervisors endorsed the statement: "My immediate bosses have a very good idea as to whether I am doing good work or not." In one company 66 per cent of the rank and file, but only 36 per cent of the supervisors expressed this opinion. In one company, 78 per cent of the rank-and-file employees, but only 42 per cent of the supervisors checked the statement: "In talking with my bosses about my job, I feel completely free to say exactly how I feel about it." The average of the percentages of supervisors and the average of the percentages of rank-and-file employees in these thirty organizations who checked this statement differed by only one point (66 and 65 per cent, respectively).

A comparison of employee- and supervisory-attitudes within a given company will sometimes point to a weakness in top management's handling of supervisors which is reflected in the way supervisors treat the employees under their direction. It usually follows (with some striking exceptions) that if the supervisory staff is unhappy, the rank and file cannot be expected to be happy. But supervisors may have favorable attitudes toward the company and their jobs, while the rank and file may have relatively unfavorable attitudes.

In addition to questions dealing with general job satisfaction, quality of supervision, advancement, pay, training, and orientation, the questionnaire may contain items about physical working conditions. Many of the specific phases of physical surroundings can be checked on through simple observation, but often bad conditions are overlooked until employee criticisms are voiced on the survey. Questions covering eating facilities, light, heat, ventilation, locker rooms, time clocks, rest rooms often reveal sources of minor dissatisfactions that affect work.

Sometimes the source may be psychological, but nevertheless important. On the recommendation of the lighting engineer all the lights in one department were changed. The new arrangement resulted in much better light for all as measured objectively. Soon thereafter a survey of employee attitudes was made in this company, and nearly every person in that department com-

plained about the light being too weak. Some employees requested specifically that the old lighting system be restored.

The plant safety engineer (in smaller plants the safety committee) should be able to check up on anything as concrete as a safety program. Yet in a manufacturing company proud of its safety record and its safety program, relatively large proportions of the employees in some departments rated the accident prevention efforts of the company as poor. They stated they had not received as much training in safety as they needed, said that the enforcement of safety rules in their departments was not strict enough, and reported that there were accident hazards in connection with their jobs which the company could eliminate. Top management was stimulated to review the company's safety program.

Most companies have certain employee benefit policies, such as group life insurance, retirement plan, sick benefit plan, vacation plan, credit union, and employees should be fully informed about these. When an employee attitude survey is being made, there is opportunity to find out whether all employees have received the desired information about these policies, and also about social security and unemployment insurance.

Many companies have used a portion of the questionnaire to get the reactions of employees to company newsletters, magazines, or other publications for employees.

Though a minimum of information about the employee should be requested in the questionnaire so as not to arouse any suspicion as to whether a blank can be identified, it is logical to ask for at least two items of information. One is sex, especially if questions regarding rest rooms are used. The second is approximate length of service, especially if questions are included that can be interpreted properly only by knowing the length of service-- such as, information provided or treatment in personnel office when first hired.

Analysis of responses by sex and by length of service, especially in the larger departments, will be found of assistance in interpreting the attitudes and opinions expressed.

Some of the most pertinent values of the survey of employee attitudes and the topics usually covered in the questionnaire have been discussed in this chapter. Other topics and other uses may suggest themselves in a review of the problems that confront management in the particular company in which the survey is made. Other uses of the data may also be found as the responses of the employees are studied and analyzed.

Chapter XVIII

THE LAST THREE FEET

CHESTER E. HARING
Committee for Economic Development

RECENTLY the theory was advanced that the last three feet of practically all our production lines is the three feet of counter space separating the prospective consumer from the retail sales clerk. Although many national manufacturers may not recognize this fact, in recent years most of them have spent considerable time and money on the education of retail dealers and their salespeople. Retail surveys can frequently supply the answer as to the effectiveness of such work.

This is only one of the purposes for which retailing research may be conducted. Sometimes dealer studies are conducted simultaneously with consumer surveys, in an attempt to determine the accuracy of each set of replies on consumer buying habits. Dealers have frequently been quizzed about their attitudes toward the distribution policies of certain manufacturers. One recent shortage item, a dealer survey showed, was the cause of negative attitudes toward the maker, since dealers thought he showed favoritism. Dealers have been quizzed on their habits, attitudes, and sales policies of certain items. Again, in a recent study, one of the purposes of the dealer study was to determine what brands of commodity were strong in each section.

Dealers have even been questioned on their opinions about the best advertising medium to reach their customers for a particular product. They have been asked to give their opinions on various types of advertising copy. When the manufacturer wanted to know how to reach the dealers through advertising, then dealers have been asked about the general and trade publications they read regularly.

However, it is probably safe to say that *most* dealer studies are concerned with retail selling methods. There is even one agency in New York which has specialized on this type of work. This firm observes and analyzes selling

methods in retail outlets, then recommends methods of improvement. It is claimed that the organization can show definitely that thousands of dollars of increased sales have resulted from putting their recommendation into practice.

Dealer studies are not always confined to dealers. The survey may be conducted among the dealer's customers, or it may be made among salespeople. In one case a department store was interested in measuring the influence of charge accounts upon their owners. A cross section of women in town was interviewed, and attitudes and buying habits relating to all department stores were measured. Toward the end of the questionnaire, there was an inquiry about the stores in which the family had accounts. Possession of accounts was then related to attitudes and buying habits. It was found that possession of an account did indeed play a tremendous role in both shopping habits and department store attitudes.

In other cases the salespeople themselves may be asked questions relating to their activities. Even more common than this is the type of study known as a "shopping survey." In this case the interviewers act as customers, with standard instructions on what procedure to follow. They make notes following completion of the contact on various elements of the treatment accorded them. The salesperson, of course, believes that the interviewer is simply another customer, in order to prevent biased results.

A DEALER SURVEY

Probably as much dealer educational work has been done by the petroleum industry as by any single industry. When you drive into a service station and the attendant greets you with a cheerful "Good morning" followed with "Shall I fill her up?" you never stop to think that this is the result of much training and educating by the oil company whose products he is selling. Most motorists would be amazed if they knew the whole story behind the average service station, whether company or privately operated. Locations are carefully and skillfully selected. Buildings are designed by good architects. Layouts are planned by well-trained engineers. Cleanliness of buildings and premises are the result of much company merchandising effort. Signs, windows, and the display of merchandise are all carefully studied. How to dress, how to greet motorists, how to sell all products and services—every detail of operation down to the final "Thank you"—all are the results of countless experiments, ending in definite suggestions and recommendations. Furthermore, all these things receive constant and close attention from the company representative who has, in turn, been trained in sound merchandising practices.

Therefore, when the Hooper-Holmes Bureau was asked by one of the large petroleum companies to conduct a study of the effectiveness of all their educational and merchandising work at service stations the opportunity of doing a worth-while dealer survey was welcomed.

Quite early came the realization that we would not succeed in gathering the *facts* desired if efforts were confined merely to questioning dealers. Surely every research man recognizes that if he asks the average dealer in men's furnishings if he suggests the purchase of a tie to every shirt customer, the answer will invariably be "Yes." Every dealer has been told to do so by all shirt and tie manufacturers. Perhaps he does so occasionally, but any survey asking such a question is bound to produce misleading results. Therefore, it was early decided to make this retail study largely a survey of observation. Accordingly, the entire questionnaire was constructed along a research technique developed many years earlier by the writer, called the Theory of Last Purchase.

The theory of last purchase. Consumers have great difficulty in answering such simple questions as: "Where do you buy your hand bags?" "What price do you pay?" "What material do you select?" etc. Obviously, most consumers do not always go to the same store every time they buy a handbag and many other articles. Nor do they always pay the same price or select the same material. This *real difficulty of answering questions in consumer research* would be obviated if questions were confined to the specific bag or other article last purchased. Then, if a proper cross section of the consumer market were obtained, final tabulations would produce a reliable picture of present buying trends. For instance, most men will have considerable difficulty in telling where they buy their ties, but none at all in answering detailed questions about the last tie they bought. Thus developed the theory of last purchase.

Accordingly, the questionnaire developed for this retail study contained exceedingly few questions to be asked of the dealer, and was made up almost entirely of places to record definite and concrete observations. While stated in the form of questions, the questions were not directed at the dealer, but were asked of our investigators.

Field covered. Naturally, since this study was conducted for one important oil company, his retail distribution determined the districts and cities to be covered. Furthermore, it was decided to confine our study to service stations whose principal business was the sale of gasoline, and eliminate all garages, parking lots, stores, and roadside stands which might be selling the client's or competing products.

Interviews were made in all states of the client's marketing territory, and since his distribution is sectional, this cannot be considered a national survey. He furnished us with the names of all cities to be visited in each division and

with the names and addresses of his stations which were to be interviewed.

Of course such a retail survey would *not be complete without learning* about the accomplishments of competition, for one of the final objectives was to compare the results of the client's work with that of his competitors. To this end the client provided names of Competitor A and Competitor B and it was decided to confine calls on competition to these two leading competing companies.

The location of each client interview in each city determined the location of all interviews with competitors. This was clearly stated to the Hooper-Holmes Bureau representatives whose instructions on this point read: *The selection of a given client station is to determine the location of the competing stations. In other words, if you cover a client station at Main and High, for example, then also cover the nearest Competitor A or Competitor B or both, for one or two competing stations. In other words, we want to cover stations which actually compete for local business and frequently you will find our client's station and both of the others in the immediate vicinity.*"

As a result of these specifications 1,267 reports on the client's stations and 2,338 reports on competing stations—3,605 in all were obtained.

Automobile requirements for interviews. Perhaps the best way to show what was required on car conditions before each interview will be to quote that part of our instructions to our Hooper-Holmes investigators:

No contact is to be made unless your car meets fully with ALL these requirements:

- a. Never have more than 5 gallons of gasoline in your tank—less will be better.
- b. Be sure your oil supply is obviously low—down one or two quarts.
- c. Before making any contact deflate the tire nearest the gasoline filler cap. Let enough air out of this tire to make it obvious that it is under-inflated, but not so low as to ruin your tire. (This can be done a block away from the station.)
- d. Be sure your windshield is dirty before making each report. Use your own ingenuity as to how this should be accomplished. In wet weather there is no problem. Perhaps a cheap atomizer or spray gun filled with dirty water will accomplish your purpose. We question the advisability of wiping it with an oily cloth for this would be too obvious. In some sections dust may be present or available.

The necessity of these requirements will be obvious as you study the questionnaire.

Format of questionnaire. The questionnaire used was printed, and consisted of four 8½ x 11 pages. This seems rather long, but actually it was unusually short. On the first page were 15 numbered questions. The first three merely identified the interview as to name, address, and petroleum company

whose products were sold. Of the remaining questions only three had to be asked, for all the others were directed at the investigator and merely required the recording of his observations.

Our client had a radio program and naturally wished to know how familiar his dealers were with it. This required a direct question to the attendant. Competing stations were asked a similar question about their company. Incidentally 91.2 per cent of our client's dealers were familiar with his program.

Our client also had a direct-mail advertising service which he sold to his dealers. Of course his records showed which stations had subscribed for this, but to check subscribers and nonsubscribers on our reports against the facts disclosed by them, required our asking a direct question.

The final question on page 1 asked the average monthly sales in gallons of gasoline. This is the kind of question that most investigators and research men fight shy of. Our instructions read: "This should present no trouble if asked in an off-hand manner. Something like this: 'About how many gallons of gas do you pump here per month?' Ask of all stations. If weekly gallonage is given, translate it into monthly by multiplying by four." In short, we knew these data are always required on report forms and all station operators know their gallonage. We assumed that when approached properly they would divulge the information needed. Our assumption was justified because 92.6 per cent of our client's stations and 95.0 per cent of competing stations gave the information wanted. Strangely enough, our final tabulations showed almost exactly the same average monthly gallonage for our client's stations and those of his leading competitors—slightly over 11,000 gallons for each.

Pages 2 and 3 of our questionnaire were made up entirely of fairly large reproductions of various display cards and posters furnished by our client, and for the most part all our field men were required to do was to indicate the presence of this material in its proper place. Page 4 consisted of four questions, all directed to our representatives, covering the use of similar display material by competing dealers.

Courtesy services. Possibly the services offered by most gasoline service stations have received more favorable comment from the motoring public than those offered by any other kind of retail dealer. Undoubtedly many manufacturers have wished that their dealers would so promptly and courteously offer similar services. Naturally, one of our questions covered the commonest of these services. It read as follows:

Check in the proper column just how your car was serviced on the following:

- a. Checked Oil
- b. Checked Water
- c. Wiped Windshield
- d. Checked Tires

To the right of these items were three columns headed: Rendered Service Without Asking, Offered to Render Service, and No Attempt at Service.

The results of these observations in 3,605 service stations (client's and competitors') was:

<i>Service</i>	<i>Without Asking</i>	<i>Offered To</i>	<i>No Attempt Made</i>
a. Checked Oil	39.2	42.9	17.9
b. Checked Water	42.9	35.2	21.9
c. Wiped Windshield	67.0	4.5	28.5
d. Checked Tires	21.8	20.7	57.5

For the entire study there was practically no difference between the service rendered by our client's stations and by those of his competitors, nor was there much difference between company-operated stations with salaried managers and dealer-operated stations. But of particular interest to our client were the variations between his several sales areas and the cities and states composing these areas, which he of course could interpret in the light of personnel and local management.

In studying the results above it should not be overlooked that our representatives laid themselves and their cars wide open for service. For instance, even though the tire nearest to the attendant when he put in the gasoline hose was seriously underinflated, 57 per cent were unnoticed. This is the measure of the task facing merchandisers to train sales people to do the few things which the manufacturer knows result in greater good will and increased sales. And it should not be forgotten, too, that these results were obtained in a field which deservedly has a high reputation.

Sales of premium merchandise. The tremendous task faced by national merchandisers in educating retail sales people is brought out by our observations of their selling and not selling premium merchandise.

Practically all service stations sell a regular and a premium (Ethyl) gasoline. Perhaps the instructions to field representatives on this point will be enlightening: "When you drive into the client's stations, stop your car between the premium and regular pumps." (Brand names were given in the actual instructions.) Similar instructions covered competing stations. Our instructions continued: "If the attendant does not open the conversation, say 'Give me five!' Begin all your observations at this point. Report EXACTLY what he says."

In nearly nine cases out of ten our men reported the attendant opened the conversation and reported his exact words. But in only four cases out of ten did he suggest which grade of gasoline the motorist (our representative) should buy. Furthermore, in approximately half of the times when the attendant did make a suggestion, he suggested the regular grade of gasoline. So once more the difficulty of training retail sales people becomes evident.

This difficulty is further emphasized by the observed results when we said "Give me five." Remembering that before each contact our men had siphoned off or drained off practically all of the gasoline in their tanks, it is startling to report that in only approximately 15 per cent of the cases did the attendant make any attempt to sell more gas than was asked for.

Rest room situation excellent. Since service stations generally have a high reputation for their rest rooms, obviously in such a study as this our client desired factual comparisons of his rest rooms with those of his competitors. If he hoped to make a better showing than competition, we had to disappoint him. The standards of both were high and final tabulations show only unimportant differences in percentages. For instance, about 80 per cent of all were observed to be in a clean condition, and over 85 per cent were fairly modern. Soap, towels, and toilet paper were present in 8 or 9 out of each 10. (Wives of Hooper-Holmes Bureau representatives accompanied most of them on this study which made the calls more casual and also more complete.)

Resulting action. Following his study of the results of this survey, our client held meetings in the field, discussed the entire merchandising effort with his field people, and took steps to improve the effectiveness of his organization. He accelerated his training courses for his own salesmen who contact station operators, and simplified some of his training programs for both salesmen and service station people alike. Special effort was put forth to strengthen his organization at those points where weaknesses were indicated by the survey.

The study has great value as a pattern for future research of a similar nature. Many manufacturers can profit from similar studies which can show them how well or poorly their educational and promotional past efforts have succeeded. Properly conducted, such studies can also point the direction to be taken in postwar sales training.

In making such studies it should be pointed out that in most cases these should be studies in observation, duplicating as completely as possible normal customer-salesperson relations. In this case it should be noted that in no instance did the service station attendant know he was being observed or interviewed or even suspect that the entire transaction was anything except entirely normal.

Retail selling actually is the last three feet of almost all production lines, and almost any sum spent upon its exploration will pay large dividends.

PART TWO

SURVEYS REPORTED TO GOVERNMENT

Chapter XIX

THE SURVEYS DIVISION OF OWI: GOVERNMENTAL USE OF RESEARCH FOR INFORMATIONAL PROBLEMS

DANIEL KATZ
Brooklyn College

THE Surveys Division of the Domestic Branch of OWI was created to serve as the eyes and ears for administrators concerned with informational policy on the home front. Its functioning can be understood most fully, therefore, in the context of the domestic operations of OWI.

The role of government-controlled propaganda on the home front during wartime vividly illustrated the contrast between totalitarian state and democratic nation. The fascist state controls all media of communication and information in order to mold the thinking of its people and to inspire and maintain morale. In our democracy, however, not only are all media for informing and influencing people in private hands but the government does not even attempt a propaganda program designed to control civilian morale in a *direct* manner. Its informational activities were restricted to acquainting people with specific war measures such as war bond drives, salvage efforts, rationing rules, and the blood donor campaign.

This limited role is an integral part of the emphasis in our democracy upon voluntary participation and genuine consent of the governed. Many necessary measures for the war effort lacked legal sanctions from the beginning of the war. There was no draft of manpower, no draft of capital, no governmental compulsion to contribute to salvage campaigns or to buy bonds. Even rationing and price control were set up with a minimum of enforcement machinery. In short, in this total war we leaned over backwards to make full participation on the home front a voluntary matter for our citizens.

One consequence has been the necessity of informing people of specific war measures and of urging them to take part. This became the main function of the Domestic Branch of OWI. In publicizing these programs, OWI did

not have the benefit of a controlled press or a controlled radio. Newspapers and radio stations were free to treat governmental messages as they desired. Moreover, the relatively insignificant budgets allotted by Congress for domestic informational activities (about \$8,000,000 in 1942-1943; \$2,500,000 in '43-'44; and \$2,250,000 in '44-'45) meant that existence of these activities depended in great measure upon voluntary efforts of radio networks, newspapers, magazines, and business and advertising firms. During 1943-1944 the business world contributed roughly \$300,000,000 in printed media space and radio time to governmental appeals channeled and directed through OWI.

The main task of the Surveys Division thus was to serve as the agency for evaluating effectiveness of the specific informational campaigns which went to make up the \$300,000,000 of space and time contributed by private industry. Its activities consisted of these types of research operations: (1) *market research*, (2) *opinion and attitude studies*, and (3) *intensive social-science research* on special problems. All three types were based upon the interviewing method.

The *market research* of the Surveys Division can in turn be broken down into three kinds of studies:

(a) *Pre-testing* the comprehensibility and appeal of informational materials on a small sample prior to their public release. For example, investigating the effects of a security of information (don't talk) program on a small audience or trying out various forms of instructions for the use of the OPA rationing booklet to see which was most readily understood.

(b) *Auditing*, or checking on, the effectiveness of the distribution of informational materials through a given medium or method. For example, finding out how many posters mailed out from Washington were actually displayed in store windows on the Main Streets of the nation during the week they were scheduled to appear.

(c) *Studying the effectiveness* of information campaigns through repeated tests of the increase in popular understanding, popular support and participation in these programs. For example, correlating changes in knowledge and understanding of inflation with the amount and kind of governmental messages from time to time.

The *opinion and attitude type* of study was conducted sometimes independently and sometimes as an integral part of a market research operation to discover the underlying attitudes, emotional resistances and conditions affecting the support of war measures. For example, how are attitudes of optimism and complacency related to participation in salvage programs, or what are the resistances to taking war jobs or the reasons for leaving war jobs? The purpose of the attitude or opinion study is to give over-all guidance to informa-

tional campaigns, to help in their planning, and to get at fundamental causes for the lack of identification with certain phases of the war effort.

Finally special problems emerge, which need more thorough and specialized study of the type ordinarily carried out by a scientific foundation. For example, during the winter of 1942-43 absenteeism in war industry was the subject of popular attention, and OWI was confronted by the problem of the role of informational policy in helping to remedy the situation. The fundamental question that had to be answered concerned the nature of absenteeism itself. Was it largely due to malingering war workers with swollen incomes, or was it a function of employing married women with home responsibilities in congested areas of bad housing and transportation? If the former analysis was correct, then OWI might want to join the newspaper crusade against the irresponsible war worker. If the latter interpretation was correct, then the solution lay in remedial measures within industry and within the community. To get an answer about the important causal factors was a matter of field interviews with key informants such as industrial personnel managers and community leaders, and with rank-and-file war workers. Moreover, only by going into the field and checking plant records was it possible to get an accurate measure of absenteeism.

Save for a few studies in the early part of its existence, the Surveys Division undertook no general-purpose research. Its work was always directed at a specific problem administrators wanted answered. During one year it operated in close relation to the OWI program managers entrusted with specific informational campaigns. Studies were initiated by the program manager who preferred, not to play by ear, but to have exact data on how his campaign was reaching the people and affecting their support of and participation in it. Occasional studies were also undertaken at the instigation of other governmental agencies when their problems related to informational policy or public opinion.

The description of one or two specific studies in terms of operational procedures will illustrate both the details of the research process and its useful applications. The two studies which may perhaps be cited as good examples are (1) the Continuing Investigation of the Effect of the Anti-Inflation Campaign and (2) the Absenteeism Survey.

THE CONTINUING STUDY OF THE EFFECT OF THE ANTI-INFLATION CAMPAIGN

At intervals of about three months, nation-wide surveys of the popular understanding of anti-inflationary measures were conducted. A basic core of questions was repeated from study to study to give comparative data on in-

creases in knowledge and understanding. The procedure for this type of study follows:

Planning phase. At the request of the campaign manager in charge of anti-inflation messages, the Surveys Division worked out a preliminary plan for a study that described the general method and type of sample, stated specific objectives, and gave a cost estimate. The preliminary plan, for example, called for a large enough national sample to permit breakdowns according to region of the country, education, economic status and occupation. It called attention to the need for a trend study since an absolute measure of knowledge at one point in time would not have much meaning. The objectives were twofold: a measure of knowledge and a measure of attitudinal support.

The campaign manager studied the preliminary plan and called attention to the desirability of checking on the medium, whether radio, newspaper, or magazine, to which people had been exposed. He was interested in finding out the relative value of approaching the people through different media. The preliminary questionnaire therefore took into account this suggestion, as well as other suggestions resulting from the conference with the campaign manager. Moreover, the objectives of the study were sharpened by the manager's specification of the kind of messages he was emphasizing in trying to reach the people.

The Pre-test. Even while the conversations with the campaign manager were going on, the pre-test had begun. The director of the study, with the help of expert home-office interviewers, was trying out the questionnaire on a small sample of people. This pre-test had two purposes: (1) To find out the actual dimensions of the problem in popular thinking and feeling and (2) to develop clear, adequate and comprehensible questions. The first function was achieved mainly through open-ended intensive interviewing in which the respondent was not limited to fixed alternative answers but was encouraged to talk freely. The second purpose was carried out by the constant reformulation of questions until interviewers felt that they were getting pertinent and meaningful answers.

Intensive interviewing showed the necessity of measuring knowledge of the anti-inflation campaign through a number of methods. The free recall method (in which the respondent was given practically no cues) underestimated the amount of his knowledge, whereas the recognition method (in which he was asked what effect rising prices would have on specific economic processes) tended to overestimate his knowledge. The final questionnaire first asked the respondent to tell what he could about the effects of inflation or rising prices, and later gave a series of recognition tests.

More important, the intensive pre-test interviewing showed the significance

of distinguishing between knowledge and understanding. Some people, for example, would give the slogans of the anti-inflation campaign without being able to apply that knowledge to related situations. They knew that keeping down wages and prices would prevent inflation, but did not understand how bond buying and heavier taxes, i.e., reducing the amount of money in circulation, was related to inflation. Questions were devised to test understanding as well as knowledge.

Putting the study into the field. After a week of pre-testing the final questionnaire was ready. Field interviewing throughout the nation was done by the National Opinion Research Center, which also assisted in the formulation of the original questionnaire and pre-testing.¹ The Center's experts also prepared instructions to its field interviewers. These included a clear statement of the objective of the study, background material for the information of the interviewer, and detailed specifications covering every question in the schedule. The specifications had both the purpose of clarifying the question for the interviewer and of standardizing his method of dealing with any problems that arose in the respondent's reactions to a question.

Quotas were assigned by the Center to interviewers, and were so drawn up as to give in the sample of five thousand respondents the right proportions of the age, sex, income, and geographical groupings of the population.

Finally, interviewers were asked, upon completion of their quotas, to give a detailed evaluation of the validity of the questionnaire, item by item.

Analysis of results. As the completed interviews arrived at the home office, answers were coded and codes constructed for the free-answer questions. In the code construction a thorough sampling of free answers was made, and these answers were classified according to their significant dimensions. These classifications formed the first crude codes, and the coding staff then put them to the empirical test of using them as categories for classifying all answers. In this process it often became necessary to revise the code until a working set of categories emerged.

Questionnaires were coded so that answers could be punched on Hollerith cards. The cards were then machine-tabulated so that cross-tabulations or correlational tables could be quickly obtained.

In analyzing results the study director first asked for straight runs on all questions, that is, a simple statement of the percentage of people giving various types of answers to each question. These straight runs would tell part of the story of the effectiveness of the anti-inflation campaign when compared

¹ The collaboration of the National Opinion Research Center was obtained on almost all studies. Throughout its existence the Surveys Division had a contractual arrangement with the Center for field work and assistance in questionnaire construction.

with a similar study in the future. Such comparison would indicate whether there was an increase in knowledge, understanding, or attitudinal support.

Straight runs, moreover, provide some basis for determining the necessary cross-tabulations. A cross-tabulation merely means a comparison of the way in which the people answering one question answered another question; for example, how many people who reported hearing anti-inflationary messages on the radio as compared with those reading it in the newspaper have an understanding of the relation between inflation and taxes. The kind of cross-tabulation to be made was also determined by the original plan of the study. For instance, since understanding of inflation is in part a matter of education, it was necessary to hold education constant in the analysis of results. This was done by dividing the respondents' cards into three groups: high school graduates or better, grammar school graduates lacking a high school diploma, and those who had not completed grammar school. Comparisons or correlations of answers were always made within these three groupings so that comparisons would not be affected by educational level.

The analysis started with the questions which gave a measure of exposure to the anti-inflationary campaign. People who reported reading or hearing governmental messages were regarded as the exposed group, in contrast to respondents who could not remember having heard or listened to such messages. The exposed and unexposed groups were compared with respect to (a) their knowledge of inflation, (b) their understanding of the role of anti-inflation measures, and (c) their attitudinal acceptance of the governmental program. In these comparisons, education was held constant as previously described so that the unexposed people with only grammar school training were contrasted with the exposed people of similar educational background.

A further refinement in the analysis was a comparison of the degree of exposure. The number of media through which the individual had been reached was compared with the extent of knowledge and understanding. In addition, effectiveness of radio could be measured in relation to effectiveness of newspaper. Still another type of analysis concerned the relation of understanding of inflation to attitudinal support of anti-inflationary measures.

Results of this analysis were then made available to the campaign manager in the form of a preliminary report.

Final report and the use of results. The preliminary report to the campaign manager indicated that the campaign had been very effective. People who had been reached by it knew and understood more about the problem of inflation than the unexposed group. This still left open the question of whether the greater knowledge among the exposed people was a function of the governmental messages, or whether it was due to a greater original interest in the government's war programs. Hence, the campaign manager wanted an addi-

tional cross-tabulation to compare the messages recalled with the type of medium to which the individual was exposed. Since different points had been stressed in the various media, further correlation between type of exposure and type of knowledge would help resolve the causal question in favor of the effectiveness of the governmental campaign. This problem emphasized the need for repeating the study to see whether additional messages would increase knowledge and understanding.

Cross-tabulations of type of message by type of medium furthermore gave the campaign manager an opportunity to see how well specific parts of his program were going. In general, the campaign manager derived the following types of help from this and subsequent studies:

(1) It was definitely indicated, as the study was repeated, that exposure was successful in increasing knowledge and understanding. The use of radio time and newspaper and magazine space for this informational program was therefore justified.

(2) The studies showed the particular parts of an anti-inflationary program which were so well understood that they no longer needed repetition. Similarly, it demonstrated those problems related to inflation which people still failed to understand, and hence, needed more information about.

(3) The surveys indicated the great importance of reaching people through more than one medium. The particular medium used was not as important as the number of messages encountered by the person. The individual who had been reached through newspaper, radio and magazine, was definitely superior to the person who had been reached through only one medium.

(4) The effect of exposure upon attitudinal support of government measures to combat inflation was less significant than it was upon knowledge and understanding. Even some of the people who understood the meaning of inflation were not always in favor of measures directly affecting the control of their own incomes. Results indicated that for thoroughgoing support of the governmental program it would be necessary to create a different psychological atmosphere—an atmosphere based upon a conviction of the equality of sacrifice and upon personally relating the sacrifices on the fighting front to the everyday activities of the housewife in her shopping and household management, or the wage earner in the disposition of his income. (Readers may recall the implementation of this notion in the car card of the soldier and the slogan: "Can they get by on my pay?")

A SURVEY OF ABSENTEEISM

The survey of the causes of absenteeism and of the methods for dealing with it illustrates another type of research conducted by the Surveys Division.

Planning phase. The planning phase was more complicated than in the inflation surveys. The OWI was interested in absenteeism as it related to informational policy. The War Production Board, the War Manpower Commission, the Army and Navy were concerned with maximum production of war materials. Accordingly, a joint committee representing all these agencies was set up to guide and to assist the conduct of the survey. The common point of agreement of this group was exact information about the incidence and causes of absenteeism. As the study progressed, the War Production Board developed a special interest in the methods of treating absenteeism.

The plan devised by the Surveys Division for the study called for two separate types of investigation: (a) The determination of the incidence and causes of absences by working out the objective relationships between the attendance record of the workers and plant conditions, community factors and personal characteristics. (b) The utilization of a special sample of the last one hundred absences in each of two plants to find out exactly why these specific people had stayed away from work on this particular occasion.

The first quantitative investigation was designed to include twenty plants representative of shipbuilding, aircraft, and ordnance. In each of the twenty plants it was planned to interview a random sample of one hundred workers by taking every *n*th name on the payroll list. At the same time, the worker's absence record for the past three months was noted to give an objective measure against which possible causal factors could be correlated. In this investigation it was planned to go beyond the interview with the representative sample of workers and to obtain all the factual information and expert knowledge about the plant and about the community. Thus, the study called for three schedules: (a) the interviewing schedule for the worker; (b) the interviewing schedule for the expert informant, whether personnel officer or plant infirmary-head or community leader; (c) a schedule for the field supervisor in charge of the study at a given plant to indicate the needed factual data about plant and community conditions. The material obtained through the latter two schedules was to furnish the background for the findings of the study. In some instances these materials would constitute findings in and of themselves.

The second part of the study conducted with absentees themselves was aimed at the reasons and values of the worker who took time off from his job.

Pre-test. The questionnaire developed for interviewing workers consisted in large part of inquiries to elicit their reactions to all phases of their working situation. Questions about absenteeism were incidental to this interview, and the respondents did not know the specific object of the study. As the pre-testing proceeded, questions were formulated to find out the worker's satisfaction with his earning situation, hours of work, safety conditions in

the plant, physical conditions of work, eating and health facilities in the plant. He was also asked about the utilization of materials and manpower in the plant, and about his community conditions of living. In addition, the interviewing schedule called for factual information about actual traveling time to and from the plant, length of time in industry, length of time in particular plant, actual present earnings, previous earnings, length of residence in the community, etc. The worker was also questioned about whether spouse was working, and if a married woman, about number of small children.

The pre-test for interviewing the absentee showed that the necessary psychological information could only be obtained through intensive, relatively open-ended interviewing. It was necessary here first to establish good rapport with the worker, and then to develop a series of probes or specific "why" questions so that the respondent would be pushed beyond his first rationalization of his absence. The pre-testing indicated that there was very little stigma attached to occasionally staying away from work. For the factory man or woman to take time out on occasion because of illness, family responsibility, or other emergencies was as natural as going to work itself. The absenteeism which constituted a "Problem" in the newspapers was not, in their thinking, related to their own daily lives. Hence, they talked freely about the whole problem of unscheduled staying away from work.

Putting the study into the field. The launching of this study was extremely difficult because each of the plants studied constituted a research investigation in itself. Field supervisors of the Surveys Division were given specific plants to cover and had the responsibility of the background interviewing with expert informants, setting up the sample, and interviewer supervision. Instructions for their many tasks had to be spelled out in advance.

Analysis of results. When the results were analyzed, significant information was obtained about

- (a) Rate of incidence of absenteeism in given types of war industry. Figures indicated that if allowance were made for new women workers with home responsibilities and for usual winter illnesses, absenteeism had not really risen above peacetime levels.
- (b) The cause of absenteeism. Causal factors lay much more in the mushrooming of war industry and the employment of women workers than in swollen wages.
- (c) Differences in absenteeism in various plants. Some plants had much better records than others and had the problem well in hand. Their specific remedial measures were noted in contrast to the plants with bad records.
- (d) Remedial treatment. The majority of plants had embarked upon no systematic method of treating the problem. In general, the more thoroughgoing the program of treatment, the more successful the plan was in combatting absenteeism.

250 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

Utilization of results. The results of the study clearly indicated that absenteeism had its roots in the natural conditions associated with the rise of war industry, and therefore the solution lay in the remedy of these conditions; for example, improved working conditions and health facilities within the plants, and nursery schools for the children of women war workers in the community. It was thus not necessary for OWI to embark upon an informational campaign designed to stigmatize the war worker as a malingerer or slacker.

The findings also pointed to the importance of the treating of absenteeism not on a national basis but in terms of the specific needs of a given plant in a particular community.

It demonstrated, moreover, the importance of worker participation in meeting the problem, and pointed to further development of labor-management committees sponsored by the War Production Board.

Chapter XX

SAMPLE SURVEYS IN CENSUS WORK

PHILIP M. HAUSER and MORRIS H. HANSEN

Bureau of the Census

Department of Commerce

THE 1940 Census in the United States completed 150 years of survey activity. This is the longest continuous experience in the use of questionnaires as a means of obtaining information from the public. However, this involved principally complete enumeration of human or material resources of the nation. In terms of sample surveys, the Bureau of the Census has had less experience. Its first modern scientific study was that undertaken in connection with the 1940 Census of Population.¹

Since 1940 the bureau has been developing sample theory and practice in the conduct of sampling surveys as a part of its regular activities.

While many forms of sampling are now used in the bureau's work, the discussion here is limited to the following applications of sampling to surveys of the population: (1) current sample surveys; (2) sampling in conjunction with the Census of Population; (3) special sample censuses; (4) sampling in conjunction with the Census of Agriculture; and (5) the proposed annual Sample Census of Population.²

Current sample surveys. Each month the Bureau of the Census takes a sample survey of the population in order to prepare a Monthly Report on the

¹This does not take into consideration a number of earlier smaller sample surveys which would not stand the test of contemporary scientific standards or the sample survey included in the 1937 National Census of Unemployment, which was designed and conducted largely by Census personnel detailed to the National Unemployment Census staff [1].

²Here the term "sample census" is used to describe a canvass of a sample of the population that is sufficiently comprehensive to provide reasonably reliable estimates of the total population of the principal areas to which it relates as well as of the proportion of the population having various characteristics, while the term "sample survey" is used to describe a more restricted canvass that is adequate to estimate averages or proportions of the population having various characteristics but not necessarily adequate to provide independent population estimates.

Labor Force. Full-time supervisors and part time enumerators are maintained in each of 68 areas. Results include estimates of the number of workers, the amount of agricultural and nonagricultural employment, and the number of unemployed, all by sex, as well as other characteristics of the various worker and nonworker groups.

The staff engaged in this work is also continuously used in making many special surveys. These may be supplemental questions to the schedule for the Monthly Report on the Labor Force, or separate sampling surveys taken for different purposes and at different times. From 1941 to 1945, war agencies made extensive use of special surveys connected with their administrative and control programs. These studies include consumer requirements, housing and housing facilities, food consumption patterns, and other subjects.

Early in the history of the Work Projects Administration there was a need for a reliable current measure of the number of unemployed. This organization, therefore, originated the Monthly Report on the Labor Force. [2] This activity was transferred to the Bureau of the Census in August, 1942, where it was further developed and maintained on a continuing basis.

The war brought still more demands for manpower information. The Selective Service System, the War Manpower Commission and other war agencies required reliable information about the current labor force. The Bureau of the Census expanded and improved the sample to meet these requirements.

The sample was modified so that reliable information could be obtained under changing war conditions. In addition, the redesigned sample aided in developing current information on changes in other population characteristics. The final sample incorporated many features of the original design but included different principles as well as new ones developed recently in the Bureau.

The sample designated a set of 68 areas (groups of 2 and sometimes 3 counties) as representative primary areas. It specified the selection of the sample of households within each area.

While this design is less efficient than a wider scattering of the sample, it was concluded that the gains from close field control and supervision were more than worth the loss. The plan emphasized quality of field work and thus reduced errors.

The scientific sampling principles utilized in selection of areas and in sampling within areas included [3] (a) the use of large and heterogeneous primary sampling units; (b) the stratification of these primary sampling units into carefully defined strata; (c) the selection of one primary unit out of each stratum for inclusion in the sample in such a manner that the probability of selection of any unit was proportionate to its 1940 population; (d) the use

of an area substratification method which provided that the sample from the selected primary unit be taken so as to represent the entire stratum from which the primary unit was drawn rather than the particular primary unit that happened to be included in the sample; (e) the subsampling of small areas within the selected primary units and of households within these small areas.

Within each county very small areas (city blocks) or small rural areas were set up. A proportion of these small areas was selected. In each of the small selected areas a complete list of all dwelling units was prepared. A subsample of the listed dwelling units was then designated for monthly coverage. The selected households are interviewed for approximately six months, after which the households in the sample are changed. The original listing of dwelling units in each area was made large enough to provide other samples of dwelling units which could be covered in subsequent periods. Between use of these subsamples, field checks were made for the purpose of listing any newly constructed dwelling units. Usually the number of small areas selected for listing were large enough to provide at least six different listings so that prelisting would last over a period of time. This not only resulted in economies, but provided a backlog of listed units which could be used for special surveys. This whole sample procedure is an application of what is frequently termed "area sampling." The method insures drawing of a sample which reflects population shifts as they occur, thus producing an unbiased sample of the population at the time the survey is made. Any changes occurring within a designated dwelling unit or in the composition of the household within a designated unit from one month to the next are reflected in the sample. Individuals are not followed from one address to another.

The interviewers speak with a responsible person in the household who can answer for himself and for other household members. If such a person is not available, callbacks are made to secure the information. One hundred per cent coverage of the listing is not obtained each month, but the number of persons not enumerated is usually below 4 per cent.³

In many sample surveys, especially in commercial research and opinion polling, the "quota" system is used, rather than "area sampling." [5] In the quota system samples are designed by estimating proportions of people in various age, sex, economic, or other groups. The interviewer is then assigned a proportionate number of calls within each of the controlled brackets. The

³ The effort to push callbacks far enough to include substantially all of the occupants of the sampled dwelling units is made because "hard to enumerate" people often have different characteristics from those more readily enumerated. For example, in families without children one is less likely to find anyone at home than in families with children, and in families where the wife is working one is less likely to find anyone at home than if only the head of the household is working. [4]

interviewer has considerable leeway in selecting respondents in this method. In addition, the method may cause a bias because the survey director lacks current information on population characteristics. In times of rapid population shift, such as in a war period, there is a danger that particularly serious biases can occur. These and other considerations led to the adoption of the unbiased sampling method previously described.

The schedules are reviewed and edited by the local supervisor before they are forwarded to Washington. A field spot check is made of a sample of the questionnaires to ensure that enumerators are "on their toes," understand instructions and are carrying them out properly. The schedules are then carefully reviewed in Washington to insure that there are no errors. Any mistakes are called to the attention of the field to improve quality of work and to establish uniformity of field work in different areas. Mechanical assignment of weights to schedules from different areas are then assigned to build up unbiased estimated totals. This results in estimates of what would have been obtained from a complete enumeration of the population. A final adjustment is applied to known population totals by age and sex groups, which improves the reliability of estimates of various subclasses of the population.

In this field organization and sample design, the Bureau has a number of advantages. It can take a quick but reliable national or local survey for continuing and changing government needs. The surveys can be obtained at a fraction of the cost of a complete census. Use of the field staff and sample for many different surveys results in a sharing of overhead costs.

Sampling in conjunction with the 1940 census of population. [6] The United States Constitution provides for the Decennial Census of Population as the basis for apportioning Congressional representation. With the emergence of this country to an urban industrial economy, additional needs for information developed from government agencies, business, industry and labor, educational and research bodies, and the general public, and the contents of the census have been expanded accordingly.

With all the expansion of the population information, it has never been fully possible to satisfy demands for additional information. In 1940 there were many problems confronting the national government. It was impossible to cover all these problems in the regular census.

Many of the inquiries could be obtained by sample surveys: inquiries about nativity of parents and mother tongue, social security coverage, veteran's status, fertility, and usual occupation of persons in the labor force.

The complete census provided a prelisting of total population, thus making possible an efficient sampling technique. By designating all persons whose names fell on specified lines of the census schedule as a special group, to be

asked additional question, a 5 per cent sample was obtained. This was a representative cross section of both persons and families.

Use of sampling methods in the 1940 Census of Population made it possible to increase the number of questions by approximately 50 per cent. This meant that the amount of information could be increased at little loss of accuracy and at considerable savings in cost. It made possible preliminary returns on the census fully eight months in advance of complete tabulation. Thus, it was possible to serve quickly the vital public need for accurate statistics about employment and age, sex and color composition of the population.

Use of sampling in the census also made it possible to increase the scope of cross-tabulations of basic, social and economic characteristics. These would have been impossible had it been necessary to run the hundreds of millions of punch cards involved.

The sample design represents an interesting example of the application of mathematical statistics to a problem involving many practical considerations and administrative difficulties.⁴

Special sample censuses. The estimate of the total population of an area through a sample census is a new project in the bureau's work. This type of activity is made feasible only because of recent advances in area sampling methods.

The War Production program caused great migratory shifts of population. These resulted in acute conditions of overcrowding in a number of "congested production areas." In April, 1943, the President appointed the Committee for Congested Production Areas. To help determine the needs of the areas for such requirements as hospital facilities, housing, food and clothing, the committee shortly wanted quick and accurate estimates of the number and characteristics of the population in these areas.

The Bureau of the Census designed and conducted sample censuses in ten designated areas,⁵ representing a population of more than ten million people. The general principles of sample design previously described were used. First, a complete listing of dwelling places by address in the smaller metropolitan areas (or an estimate based on a listing of dwelling places within a sample of city blocks within larger metropolitan areas) was obtained. The average population per dwelling was next determined. The estimate of total population of the area was then obtained by multiplying the number of dwelling places by the estimated average population per unit. [8]

⁴Information about the actual technique devised for the selection of the sample and its enumeration, the manner in which bias arising from the use of a line schedule was minimized, the type of schedule used, and other details have been published elsewhere. [7]

⁵These areas included the Puget Sound, Charleston, S.C., Detroit-Willow Run, Hampton Roads, Los Angeles, Mobile, Muskegon, Portland-Vancouver, San Diego, and the San Francisco Bay areas.

The schedule required a listing of all persons in the individual household. It included questions about the relationship of the respondent to the head of the household, sex, age and marital status. Inquiries about the labor force status of the population were included. In order to determine the number and type of new residents in the areas, a question about the place of residence of each person on April 1, 1940, was included. The question on farm residence in 1940 made it possible to estimate separately the number of new residents who came from farms.

The sampling errors of the population estimates from these sample censuses were computed. The average relative sampling error⁶ in most of the areas was about .8 per cent, while the maximum relative sampling error was about 2.4 per cent.

Within seven months after authorization of these sample censuses, a preliminary estimate of all ten areas was available. [9] More detailed tabulations of population characteristics were available shortly thereafter. [10]

Sampling in conjunction with the Census of Agriculture. The scope of information suggested for the 1945 Census of Agriculture simply does not permit inclusion of all the proposals by government agencies and private interests. By asking additional questions of a selected sample of this census, most of these subjects can be included. This technique will permit more than a 100 per cent increase in the amount of information collected in the 1945 Census of Agriculture.

Some of the additional questions such as those about forest products and commercial fertilizers represent new subjects which can be included in the census. Others, such as those about age and sex of cattle, will permit greater detail about subjects carried in the complete census. The group of inquiries about farm population and farm labor (which provides for the listing of all members of all farm households) is of considerable importance.

The construction of the sample surveys to be made here is different from those used in the 1940 Census of Population. The latter was a sample drawn from the complete listing of the population. The present design is an area sample which can be used independent of the complete farm listing. This sample (the "master sample for agriculture") consists of small tracts of land, usually about one square mile in size, averaging four or five farmsteads per unit.⁷

⁶ The average relative sampling error is the standard error of the estimated total as a percentage of the total, or, in other words, is the coefficient of variation of the estimated total, in percent. The maximum relative sampling error is 3 times the coefficient of variation, in percent. Thus it can be stated with a high degree of confidence that the estimated total does not differ from the total that would have resulted from a complete enumeration by more than the indicated maximum relative sampling error.

⁷ This design is the result of collaboration of the Bureau of Agricultural Economics, the Iowa State College, and the Bureau of the Census.

This sample will yield reasonably accurate estimates of agricultural production. It will give inventories of principal items, farm characteristics and farm population data. It will produce such figures, not only for the nation (and its regions and divisions) but also for the states.

Since it is essential in this type of sample design to identify properly the areas used, good maps and area descriptions are a prerequisite. Aerial photographs were found particularly helpful.

In addition to the expansion of subjects permitted by this sampling, the technique will give preliminary results of census statistics. Both the Bureau of Agricultural Economics and the Bureau of the Census plan to use this sample for current surveys and sample censuses of farms or farm population.

Proposed annual sample census of population. [11] This proposed sampling study is one of the most significant being planned by the bureau. Its most important single objective is the preparation of population estimates for states and large cities for years following the census.

By adding total births and immigration and subtracting deaths and emigration, an accurate estimate of total population for the country is relatively easy to make. It is not so simple to estimate the population of individual states, cities or counties, especially during periods of rapid population shift. It was possible to estimate population over states and counties subsequent to the 1940 Census only because of national registrations for war ration books. Unless there is further rationing registration, it will be impossible (under present authorized programs) to obtain any area population figures until the 1950 census of population. The sample survey technique offers the only satisfactory solution for the pressing demands about such figures made by various government and private business organizations.

A sample census of population has been designed to be taken in the fall, at least during the present period of rapid change.⁸

Questions to be included are of three main types: General population, labor force, and housing. Those relating to population will include such items as age, sex, color, marital status, relationship to family head, and migration. Those about the labor force will include the work status and perhaps questions on occupation and industry. Those concerned with housing will probably include rent or value of dwelling units, type and age of structure, occupancy status and number of rooms.

An annual sample study of this type would result in certain advantages. The scope of the work could be greatly expanded over a period of time by

⁸ Authorization and funds for the first of these annual sample censuses has been requested for the fiscal year 1947. The date for the first sample census is tentatively scheduled as October, 1947.

rotating the questions from year to year. This rotation would reduce the schedule and work load of the Decennial Census.

The sample construction would combine elements of the designs of the current urban and rural nonfarm area studies, and the master sample for farm areas. Somewhat detailed statistics would result for cities (at least those with one hundred thousand or more population) and metropolitan districts, states, regions and the total United States. Estimates for urban, rural nonfarm and rural farm population and for city size groups (for smaller cities) would be available. Such statistics would be available from a sample averaging approximately 4 per cent of the United States.

Conclusion. The sampling survey has greatly expanded the scope of normal census operations without corresponding increase in the cost of such operations to taxpayers. The sample survey has made possible current statistics on population, labor force, housing, and other items. In the short period during which sampling survey techniques have been used by the Census Bureau great developments have occurred in theory, practice, and application of sample survey methods. These developments and their results may well be regarded as merely the beginning of even greater advances which lie ahead.

REFERENCES

1. See Dedrick, Calvert L., and Hansen, Morris H. *Final Report on Total and Partial Unemployment*, Vol. IV. The Enumerative Check Census, U.S. Government Printing Office (1938).
2. This project originated under the direction of Howard B. Myers and John Webb. See Frankel, Lester R., and Stock, J. Stevens. On the sample survey of unemployment. *Journal of the American Statistical Association*, 37, (1942), 77-80.
3. These principles are described in detail in Hansen, Morris H., and Hurwitz, William N. On the Theory of Sampling from Finite Populations, *Annals of Mathematical Statistics*, 14, No. 4 (December, 1943) 333-362. For a less technical description see a paper by the same authors: A New Sample of the Population. Bureau of the Census, October, 1944.
4. See Hilgard, Ernest R., and Payne, Stanley L. Those not at Home: Riddle for Pollsters. *Public Opinion Quarterly*, 8, 254-261.
5. See Hauser, Philip M., and Hansen, Morris H. On Sampling in Market Surveys, *Journal of Marketing*, July, 1944.
6. For a more detailed discussion of this application see Hauser, Philip M. The Use of Sampling in the Census. *Journal of the American Statistical Association*, 36, (1941), 369-375.
7. Stephan, Frederick F., Deming, W. Edwards, and Hansen, Morris H. The Sampling Procedure of the 1940 Population Census. *Journal of the American Statistical Association*, 35, (1940), 615-630. For a more detailed description see Stephan, Frederick F., and Deming, W. Edwards. On the Sampling Methods in the 1940 Population Census. Bureau of the Census, March, 1941.
8. For more detailed descriptions see Brunsman, Howard G. The Sample Census of Congested Production Areas. *Journal of the American Statistical Association*, 39, No. 227, 1944.
9. See Census release, Series CA, Nos. 1-10.
10. See Census release, Series CA-1, CA-2, and CA-3.
11. See Hauser, Philip M. Proposed Annual Sample Census of Population. *Journal of the American Statistical Association*, 37, (1942), 81-88.

Chapter XXI

SURVEYS OF CONSUMER REQUIREMENTS

CHARLES E. NOYES and ERNEST R. HILGARD
War Production Board¹

THE war did not bring many serious consumer shortages in the United States prior to the spring of 1943. Exceptions were items whose production had been limited or prohibited by the War Production Board, such as mechanical pencils and mechanical refrigerators, and those which were rationed, such as shoes and canned goods. Up to 1943 the Office of Civilian Supply had frequently been in the paradoxical position of devoting more attention to things consumers could get along without in order to free war production than to the protection of consumer needs.

When the Office of Civilian Supply became the Office of Civilian Requirements in May, 1943, this emphasis began to shift rapidly. An increasing volume of protest mail, complaints forwarded through members of Congress, and reports from manufacturers and retailers showed that consumers were beginning to feel a genuine constriction in a number of fields. Farmers demanded ammunition for pest control, women wanted bobbie pins and elastic, and more seriously, there were evidences of shortages of children's clothing and work clothing. Unfortunately, most of the information about these shortages was spasmodic and inconclusive. Confronted with definite statements of military needs to the WPB Requirements Committee which had the job of dividing up scarce materials, OCR found it difficult to make a strong case for civilian needs.

It was agreed that the level of civilian supply in fields of scarcity should be the minimum necessary to maintain health and morale and working efficiency; but no one knew just where that level was or what quantities of goods were needed to support it. While considerable work had been done to justify the OCR claim for civilian goods on the basis of historical records

¹ Mr. Noyes is now associated with Press Research, Inc., and Dr. Hilgard has returned to Stanford University.

of production and consumption, the applicability of prewar figures to war conditions could not be conclusively demonstrated. Moreover, little information was available as to the extent to which retail and household inventories had been built up in 1941 and the early months of 1942. The first problem was to find out whether the reported shortages were real, or the product of chronic complainers and special interest groups.

To supply this information a Civilian Relations Division (later called Civilian Surveys Division) was established in the Office of Civilian Requirements in the early summer of 1943. As a first step, eight or ten trained interviewers were sent around the country to talk to workers in war plants and to housewives in their homes and women's organization meetings, and to the Home Economics groups sponsored by the Department of Agriculture. Several special consultants, mostly professors of marketing, were asked to supply information about shortages in their regions.

Reports from these interviewers left no doubt about the seriousness of some of the shortages. Food was being lost through lack of pressure cookers, jar tops, and other home canning equipment. Many women all over the country reported real difficulties in obtaining suitable clothing for their children and for their working sons and husbands. Especially severe at that time were the shortages of diapers and boys' overalls.

The reports were carefully tabulated by a card index system which showed what items were most frequently reported, the regions where shortages were most severe, and the reactions of civilians unable to buy what they wanted. This index was supplemented by a tabular analysis of the voluminous mail received in the Office of Civilian Requirements. By September, a list had been built up showing quite clearly the items to which OCR should devote its greatest attention. To a considerable extent, naturally, this list merely confirmed what had already been ascertained from trade and other informed sources. It did add a considerable number of new items which had not been previously considered to be in seriously short supply. It led to immediate action in severe cases, such as an increase in the supply of civilian ammunition, and some distribution of garbage pails from Army stocks.

But soon the very existence of this information created a demand for more. It was impossible to determine the extent of shortages from random interviews and reports, or even to be sure which shortages were most severe and causing the greatest hardship. It became apparent that what was needed for sound administrative decisions was an accurate nationwide survey.

FIRST SURVEY OF CONSUMER REQUIREMENTS

The use of a sample survey as a basis for specific administrative action by a government agency was a comparatively new proposal. Sample surveys had

been used by various government departments for a number of years to obtain general information on consumers' income and expenditures, and other data providing a background for the formation of policy. The Bureau of Agricultural Economics in the Department of Agriculture had done some very successful work in studying both the needs and attitudes of farmers and farm families. The Office of War Information had made a number of confidential surveys of civilian attitudes toward the war and its problems. The Bureau of the Census had used sampling techniques to expedite the availability of population statistics. But none of these uses was quite the same as the effort to obtain information as a basis for writing and changing specific regulations dealing with the production and distribution of goods.

In planning for the First Survey of Consumer Requirements, advice was obtained from leading experts in the sampling field, both inside and outside the government. In this and later national surveys the Civilian Relations Division obtained suggestions and criticism from an Advisory Committee which contributed its very valuable services.²

Fortunately, a suitable national sample had recently been developed by the Bureau of the Census, primarily for use in the Monthly Report on the Labor Force. The master sample consisted of listings of dwelling units in representative areas within 68 counties and pairs of counties so distributed as to be representative of the whole civilian population of the United States. The use of an area sample of this type was particularly advantageous for the purposes of OCR because it provided an accurate reflection of changes in consumer experience related to in- and outmigration. It was decided that a sample of about 5,000 households would be sufficient for the purposes of the first survey. The four major geographical regions, and size of community, were controlled.

The schedule for the first survey began with the following general question to introduce the subject and elicit a free answer: "The War Production Board is interested in the trouble people are having in getting things. What are some of the shortages that have bothered you most?" Answers to this question put the shortages with which OCR was specifically concerned in perspective with relation to others. Elastic tape, the top item for which OCR had administrative responsibility, was mentioned 201 times, as compared with 489 mentions of butter as a bothersome shortage.

The remainder of the schedule covered a specific list of 115 items of goods and services selected primarily from the known shortages which had been

² The members of this committee were: Dr. Theodore H. Brown, Mr. Archibald M. Crossley, Mr. Cornelius DuBois, Dr. Raymond Franzen, Dr. George Gallup, Dr. Paul Lazarsfeld, Mr. Elmo Roper, and Dr. S. S. Wilks. Mr. Austin Grimshaw, the first director of the division, served as a consultant after December, 1943.

262 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

reported by *Civilian Relations Division* interviewers and consultants, and some which were added at the special request of the OCR Operating Divisions. The basic questions covering these items in the schedule for the first survey were as follows:

- (1) "Did you buy any of the items on the list during the last two or three months?"
- (2) (For each item bought) "Was it satisfactory? If not, in what way was it not satisfactory?" (Response recorded verbatim)
- (3) "Of the items on the list that you did not buy during the last two or three months, did you try to buy any?"
- (4) (For each such item) "Why didn't you buy it?"
 - a. Price too high.
 - b. Inferior quality.
 - c. Could not find type, kind, size.
 - d. Could not find any at all.
- (5) "How much difference has it made?" (Respondent was asked to select one of four responses):
 - a. It's made no difference at all.
 - b. It's been a little inconvenient.
 - c. It's been a real inconvenience, but no hardship.
 - d. It's been a real hardship.
- (6) (If applicable) "In what way has it been a real inconvenience or a real hardship?"

Within the first survey, the time reference was shifted from two or three months to a year for durable goods and appliances, and questions were asked about presence of the items in the household, their condition, length of ownership, and difficulties in obtaining repair services and parts.

The schedule of the First Survey also contained questions about difficulties with laundry services, shoe repair, etc., and entries covering a considerable amount of factual data about the respondent household.

The decision to use the Bureau of the Census sample was not reached until October. A questionnaire was pre-tested by the Special Surveys Division of the Office of War Information and found to be, with slight modifications, feasible so far as interviewing was concerned. Interviewing on the national sample was conducted the week of November 7-13, and a general summary report on the results was printed and distributed by December 15. Table 1 is reproduced from this report. This was followed by individual reports for each item covered, summarizing in detail all of the information obtained, for the benefit of the product specialists in the Office of Civilian Requirements' Operating Division, who used the data in the preparation of OCR programs. Special

TABLE 1
ITEMS IN SHORT SUPPLY

FIRST SURVEY OF CONSUMER REQUIREMENTS, NOVEMBER, 1943

<i>By Free Mention</i>			<i>By Hardship Rating</i>			<i>By Inability to Buy</i>		
<i>Rank</i>	<i>Item</i>	<i>Mentioned In 1000</i>	<i>Rank</i>	<i>Item</i>	<i>Number Reporting Real Inconvenience or Hardship In 1000</i>	<i>Rank</i>	<i>Item</i>	<i>Number In 1000</i>
1.	Elastic tape	41	1.	Alarm clock	118	1.	Elastic tape	214
2.	Wash tub	31	2.	Elastic tape	97	2.	Alarm clock	188
3.	Sheets	27	3.	Wash tub	59	3.	Wash tub	101
4.	Pail or bucket	23	4.	Sheets	39	4.	Sheets	84
5.	Dress fabrics	20	5.	Boys' overalls	37	5.	Pail or bucket	73
6.	Alarm clock	19	6.	Iron cord	35	6.	Iron cord	64
7.	Electric iron	17	7.	Pail or bucket	33	7.	Clothespin	60
8.	Boys' overalls*	17	8.	Electric iron	30	8.	Boys' overalls	59
9.	Washing machine	16	9.	Washing machine	29	9.	Radio tubes	56
10.	Tableware (spoons, etc.)	12	10.	Women's girdle	24	10.	Pots and pans under 10 qts.	56
11.	Clothespins	12	11.5	Ammunition†	23	11.	Tableware	53
12.	Pots and pans, under 10 qts.	12	11.5	Clothespins	23	12.	Electric iron	50
13.	Girdle	12	13.	Radio tubes	22	13.	Ammunition†	49
14.	Iron cord	11	14.	Pots and pans, under 10 qts.	22	14.	Women's girdles	44
15.	Radio tubes	10	15.	Tableware	19	15.	Washing machine	41
16.	Babies' waterproof pants	9	16.	Scissors	17	16.	Women's panties	39
17.	Woven wire fence†	9	17.	Women's panties	16	17.	Zippers	38
18.	Ammunition†	9	18.	Babies' waterproof pants	15	18.	Scissors	34

*Children's clothing mentioned 114 times.

†From farm sample of 772 respondents; other items based on 4,935 replies.

reports were also issued covering information obtained from free answers, describing regional inequities in distribution, and summarizing successful and unsuccessful buying experience for all items.

The promptness with which results became available was made possible by machine tabulation in the Bureau of the Census and the combined efforts of the Civilian Relations Division staff.

FARM SUPPLIES SURVEY

The responsibility for supplying the major needs of farmers for machinery and equipment had been transferred from the War Production Board to the War Food Administration. However, the Office of Civilian Requirements continued to set up programs for the production of a considerable number of items necessary for farm production. To determine the state of supply and the need for such items a Farm Supplies Survey was carried out in February, 1944. At the request of the War Food Administration, which participated in sponsoring this survey, questions were also included about repair parts for farm machinery. Like the First Survey of Consumer Requirements, this survey was conducted in the field by the Bureau of the Census, using the farm areas of the same master sample. There were 2,678 interviews with farm operators.

Profiting by experience with the administrative use of the First Survey, the Farm Supplies Survey was designed to obtain more exact quantitative information about farmers' attempted purchases, success in buying, and anticipated needs for 1944. It was also possible to obtain a close check not only on the adequacy of the sample, but on the success with which information was obtained by interviews, through comparison of the survey results with completely independent estimates of the number of machines on farms in the United States. For example, the survey yielded a national estimate of 1,981,000 tractors on farms when the number reported in the survey was projected for the total number of farms; the January, 1944, estimate by the Bureau of Agricultural Economics, based on the 1940 Census and sales and obsolescence in the intervening years was 1,937,000 tractors.

One of the interesting and provocative results obtained from this survey was the fact that for a number of items farmers estimated their needs in 1944 as less than the quantity they had attempted to buy in 1943. This was generally true only of items which had been in fairly adequate supply, and occurred most often with respect to such things as harness and tools, the need for which is not anticipated until the ones in use wear out or get lost. On the whole, however, anticipated 1944 needs were of the same order of size as attempts to buy in 1943.

The hardship rating question used in the First Survey was modified in the

Farm Supplies Survey to include as the most extreme degree of hardship only actual interference with farm production resulting from shortages. Results for a few of the items are given in Table 2.

The replies to the hardship question appeared to be thoroughly reasonable and realistic. Only 10 per cent of the farmers who were able to buy less rope than they needed reported that it cut down the quality or quantity of their farm production, while about 45 per cent of those who had trouble getting parts or service for corn planters said that it interfered with farm production. To put it another way, flashlight batteries were reported as a shortage by more farmers than was any other item, but they ranked 38th among 43 items in terms of hardship and interference with production. At the other extreme, cream separators ranked 27th in the number of farmers affected by a shortage, but first in interference in farm production when they were not obtained.

In this, as in other surveys, it was impossible to pick one item as the most serious shortage, since different items ranked highest in the number of farmers affected, the percent of demand met, and the resulting hardship.

The survey revealed many serious shortages. Of course, the items had been selected because most of them were known to be in fairly short supply, but it is nonetheless significant to find that less than 75 per cent of the total demand had been met for 30 of the 43 items, and there were five items of which farmers had obtained less than one-half of what they had tried to buy. Realism in reporting shortages was indicated by the fact that the stated unit shortage of pressure lanterns was only three per cent of the total demand, while a few other items were reported to be in reasonably plentiful supply.

TABLE 2
SHORTAGES AND RESULTING DIFFICULTIES FOR FARMERS:
SURVEY OF FARM SUPPLIES, FEBRUARY, 1944

Item	Percent of Demand Not Satisfied	Result of Inability to Buy (as percent of those in the market)		
		Hardship, but No Interference With Production	Interfered With Production	Total Hardship and Interference
Flashlight batteries	52	39	4	43
Cream separators	37	34	48	82
Floor brooders	37	21	39	60
Harness	8	18	38	56
Implement tires (other than tractor tires)	4	51	35	86

This survey yielded some highly important information which had never before been available in any form. Of special interest was the amount of lumber used on farms, which proved to be of great importance a little later in the

year when the War Production Board found it necessary to place drastic restrictions on the use of lumber, and to institute a virtual allocation system for most of the lumber available.

SECOND SURVEY OF CONSUMER REQUIREMENTS

Distribution of detailed reports on the First Survey of Consumer Requirements almost immediately brought a request from the Programming Divisions of the Office of Civilian Requirements for more detailed quantitative information. The First Survey asked whether the families had tried to buy the various items and whether they had been successful; it did not ask how *many* they had tried to buy, or what sizes, types, or materials were in demand; or what price lines showed the severest shortages. Once the usefulness and accuracy of survey results had been demonstrated to the product specialists, they immediately asked for more specific information to be used in the development of individual item programs.

As a claimant agency representing the civilian consumer needs before the WPB Requirements Committee, it was necessary for the Office of Civilian Requirements to determine quantitatively the need for hundreds of items from garbage pails to diapers. The product specialists then translated these needs into quantities of raw materials and requested Requirements Committee approval of the resulting programs. Before the application of survey technique to this problem, programs were developed principally on the basis of prewar production and consumption figures. But there was clearly a need for information on the changed pattern of demand under war conditions to justify claims which had to compete with requirements for the military services, for agriculture, export, transportation, and other essential uses.

To improve the information available for this purpose, it was decided to attempt a much more precise survey of consumer demand and consumer needs in the spring of 1944. In order to avoid a lengthy and tiring interview, the items covered in the First Survey were divided into two parts: one, covering textile and miscellaneous household items, became the Second Survey of Consumer Requirements; and the other, concerned principally with durable goods and metal household items, became the Third Survey. The Bureau of the Census conducted enumeration and machine tabulation, with a similar but separate sample.

For 72 items of clothing and textiles, the Second Survey asked:

- (1) How many the respondent had bought between January 1, and the week of March 18.
- (2) How many were received as gifts.
- (3) The price of the last one purchased.

- (4) Whether the last purchase was satisfactory, and, if not, why?
- (5) How many the respondents had tried to buy in addition to those obtained.
- (6) Reasons for not buying, including price, quality, and inability to find any.
- (7) How much difference it made when the article was not obtained.
- (8) How many the respondent bought, received as gifts, and tried to buy unsuccessfully in the year 1943.
- (9) How much difference the shortage had made in 1943.

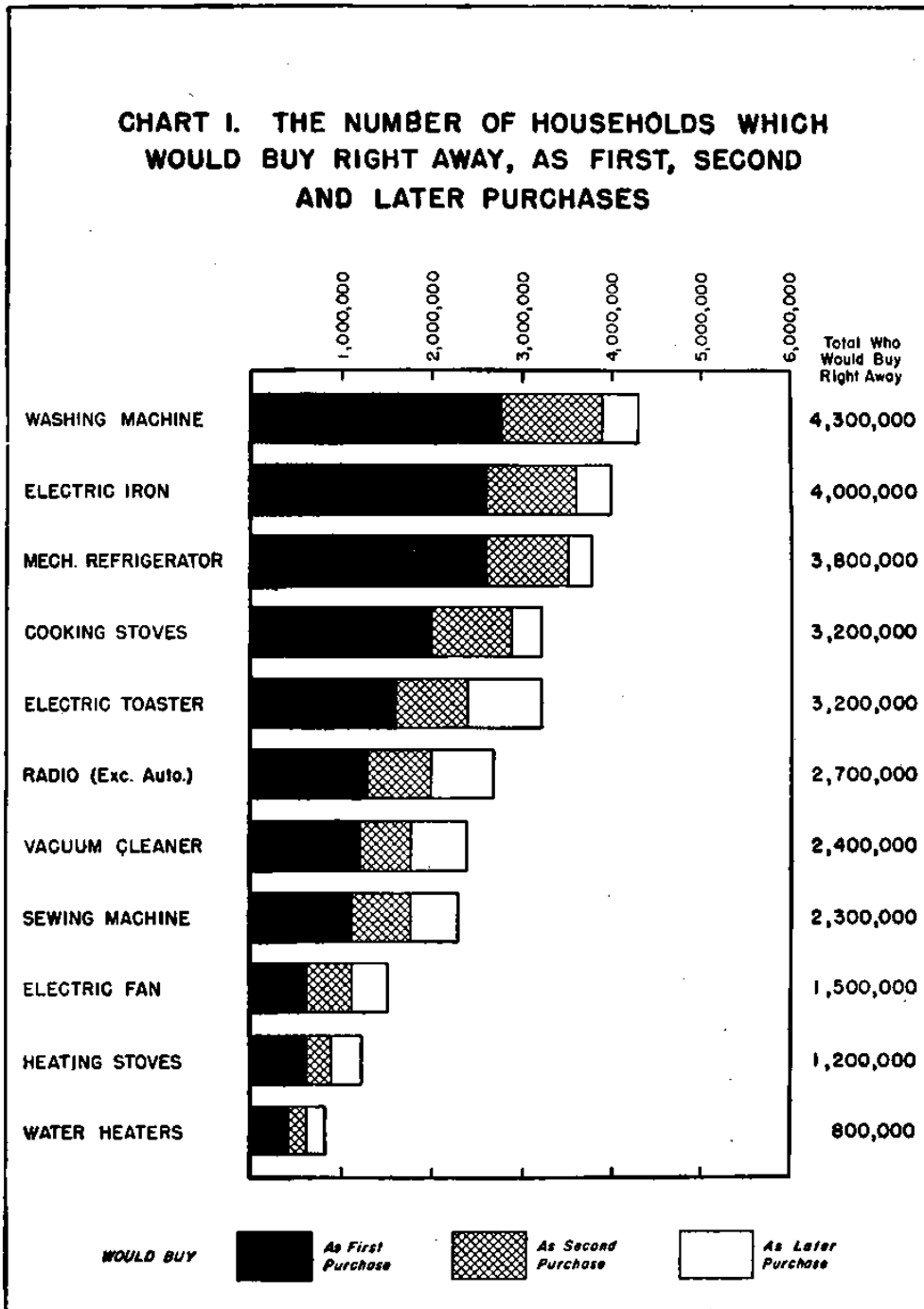
These questions were asked in relation to purchases by or for each member of the household, whose age had previously been ascertained, so the answer could be tabulated by size of garments where appropriate. This survey also provided opportunity for a free answer, and miscellaneous information about the household's income, ownership of home, etc., for purposes of cross-tabulation.

In planning the survey, and in pretesting it, it was recognized that these questions put a considerable strain upon the average respondent's memory. When the results were tabulated, it seemed likely that respondents were able to remember their most recent purchases better than those made in the previous year, since purchases reported for the first three months of 1944 averaged more nearly one-third than one-quarter of the purchases reported for 1943. On the basis of Federal Reserve Board figures for department store sales, with allowance for price changes, one-quarter would probably have been more accurate.

Nevertheless, the reported purchases for 1943, when projected to the number of individuals in the whole population of various age and sex groups, were not too different from the available production figures for the types of garments covered. (Some of the results are given in Table 3.) While it could not be claimed that the survey results were accurate within 5 or even 10 per cent, their consistency with other information rendered them usable both as a broad measure of the shifting patterns of consumer demand under war conditions, and of the degree of shortage which existed in 1943 and the first quarter of 1944 for various textile and apparel items. They were accordingly accepted as evidence to be used in modifying textile and apparel programs which had been based on prewar consumption and production figures.

Using conversion figures supplied by the War Production Board Textile Bureau, the current demand and shortage of items covered were presented in one special report in yards of fabric. This gave a very useful measure of the degree to which wartime demand could be satisfied with the amount of cloth available for civilian use when the textile shortage became more severe later in the year. Interpreted in relation to the respondent's hardship ratings, these

CHART I. THE NUMBER OF HOUSEHOLDS WHICH WOULD BUY RIGHT AWAY, AS FIRST, SECOND AND LATER PURCHASES



figures also considerably influenced administrative decisions and WPB orders directing the available fabrics into the most essential uses.

The price data collected in the survey were used to develop the programs for a number of garments by specific price lines which were incorporated in WPB apparel orders. These data were also furnished to the Office of Price Administration for use in determining the extent to which the various items were available at different established ceiling prices.

THIRD SURVEY OF CONSUMER REQUIREMENTS

Like the Second Survey, the Third was designed to obtain information about the number of items consumers had bought or attempted to buy, as well as on the existence and severity of shortages for 30 miscellaneous household items, such as washtubs, pots and pans, alarm clocks, and radio tubes. The respondent was asked whether she had bought or tried to buy any between January 1 and the week of April 18, 1944, when the interviewing was carried out. For durable goods and appliances, the time interval was the previous year.

The Third Survey also attempted to ascertain the market for appliances and for houses. Since the Office of Civilian Requirements was concerned with wartime rather than postwar demand, the pertinent question on appliances was phrased: "Would you buy a new one right away if there were plenty of everything in the stores?" Replies are summarized in Chart 1.

This question was designed to remove the patriotic disinclination to buy when things were scarce as a result of the war, but at the same time to avoid some of the uncertainty and wishful thinking which might affect the response to a question about postwar buying intentions. Its principal purpose from OCR's point of view was to ascertain the relative demand for various appliances if it should become possible to resume limited production after the fall of Germany, but before the end of the war with Japan. On the same principle of reference, the question on housing was: "Would you buy or build a new house right away if there were plenty of material and labor available?"

Respondents who gave an affirmative answer to questions about appliances or housing were then questioned about the seriousness of their buying intentions, the reason for needing each item in the case of appliances, and possible reservations which might keep them from buying right away. If they expressed an intention of buying more than one appliance, they were asked to indicate first and second choices.

Analysis of the responses to all of these questions made it possible to indicate the probable demand for appliances and new houses under varying conditions of optimism or pessimism about the future. [1] It was clearly recognized

270 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

and stated, of course, that these results indicated demand in the absence of advertising and selling pressure, attractive new models, or the desire to keep up with the Joneses.

At about the time when the results of the Third Survey became available, the Office of Civilian Requirements adopted a plan for programming durable goods at three levels: (1) a minimum essential level which would require *distribution control*; (2) a "squeeze" level which would supply most people who took the trouble to shop around actively; (3) the level of *unrestricted demand*. While as a practical matter it was necessary to consider the availability of materials and production facilities in setting the first two levels, results of the Third Survey proved very useful in keeping the figures in reasonable relationship with the existing market.

Another important use of the Third Survey results was to check the adequacy of measures already taken to increase the supply of such essential items as pails and washtubs, and to indicate which of the other items deserved special attention from the OCR Programming Divisions. Because item demand was reported, as well as households in the market, the figures were more convincing than those from the First Survey in justifying OCR's claims for materials before the Requirements Committee.

SURVEYS OF RETAIL STORES

The three Consumer Surveys and the Farm Supply Survey provided the basic information needed by OCR as to quantity of demand and degree of shortage for the *whole* country, for the four major geographic regions, and for different income groups and types of commodities. They had the disadvantage, however, of being comparatively expensive and they took a considerable length of time between the first stage of planning and the completion of item reports.

In order to maintain a more up-to-date check on supplies of items known to be short, the Civilian Relations Division arranged, beginning in November, 1943, to obtain periodic reports on the availability of consumer goods from eight large mail-order, chain, and department stores. These retail supply surveys, which were conducted every two or three months, used the following basic questions:

1. Number of weeks' supply on hand.³
2. Number of weeks' supply on hand in the same month, prewar.
3. Previous month's sales versus same month in 1942 or 1943, in units.
4. Estimate of per cent of current demand met.

³ In later surveys, stores were also asked how many days they had been out of stock during the previous month.

Some additional questions, such as the number of weeks' supply on order, were tried in the First Retail Survey, but the results were not found sufficiently useful to justify repeating them. The question on per cent of current demand met aroused considerable controversy from friendly critics within the government, but it was generally found that experienced store buyers had little difficulty in giving a definite estimate. This question was a necessary part of the survey, since there were a good many items such as facial tissue for which a store might meet a considerable part of its normal demand without being able to keep any supply on hand. If any question arose, the interviewers explained that "demand" in this question meant the estimated quantity the individual store could sell if both that store and its competitors had ample stocks.

These retail surveys could not provide information about the quantities of various items needed, but they did give a tentative check on the increasing or decreasing severity of shortages and made it possible to take prompt action in some acute cases. The information proved particularly valuable in relation to textile and apparel items when the basic pattern of consumer demand had been established by the Second Survey and OCR was experimenting with a number of different orders designed to increase the supply of the most essential textiles and clothing.

When the chairman of the War Production Board announced in September, 1944, that there would be no further use of specific programs for the production of civilian items after victory in Europe (with the possible exception of some items of clothing), it became apparent that there would be little need for additional large-scale surveys. Accordingly, plans were made to dissolve the Civilian Relations Division during the ensuing three months. However, it was essential to provide some mechanism by which additional information on the supply of consumers' goods could be obtained quickly if needed. Fortunately, the Bureau of the Census had already prepared plans for the development of a national sample of retail stores, and arrangements were made to speed the construction of a partial sample so that it would be available when needed for OCR use.

THE ACCURACY OF THESE CONSUMER SURVEYS

Because of the close tie-in between the survey results and administrative decisions within the OCR, it is probable that the detailed quantitative results of these surveys have been scrutinized and checked more thoroughly than in the case of most similar research ventures.

The ultimate consumer in most cases has a regular place of residence, so that a carefully designed household sample is ideal for obtaining an estimate

272 HOW TO CONDUCT CONSUMER AND OPINION RESEARCH

of the presence of goods in homes, and of the market for them. Our distribution system is a complicated one. People purchase not only in their neighborhoods and in nearby metropolitan centers, but buy their goods from mail-order houses, and through private ads in the for sale columns of the paper; some items, such as clothing, they often receive as gifts. The final flow of goods to the ultimate consumer can be studied only by reaching the consumer himself.

That considerable success was achieved in the consumer surveys and in the farm surveys in reaching the consumers of the goods being studied is evident through comparing the results of the interviews with independent national

TABLE 3
SURVEY ESTIMATES COMPARED WITH INDEPENDENT NATIONAL ESTIMATES

	<i>Estimate from Survey</i>	<i>Independent Estimate</i>
<i>First Survey of Consumer Requirements (November 1943)</i>		
Homes with electric irons	29,200,000	26,000,000 ^a
Homes with washing machines	17,700,000	18,400,000 ^a
Homes with refrigerators	20,600,000	19,900,000 ^a
<i>Farm Supply Survey (February 1944)</i>		
Motor trucks on farms	1,300,000	1,100,000 ^b
Mowing machines on farms	2,700,000	2,800,000 ^b
Cream separators on farms	1,700,000	1,800,000 ^b
<i>Second Survey of Consumer Requirements (March 1944)</i>		
Annual purchases compared with annual production (1943):		
Men's wool or part-wool suits	18,000,000	16,800,000 ^c
Boys' wool or part-wool suits	3,800,000	3,600,000 ^c
Men's and boys' shirts	191,000,000	182,000,000 ^c
Men's and boys' overalls, trousers, work pants	108,000,000	106,100,000 ^c
<i>Third Survey of Consumer Requirements (April 1944)</i>		
Quarterly purchases compared with reported shipments (First quarter purchases 1944):		
Radio tubes	3,800,000	4,200,000 ^d
Extension cords	1,800,000	2,000,000 ^d
Washtubs	1,500,000	1,100,000 ^d
Window screening (Sq. yrds.)	7,700,000	7,700,000 ^d
Scissors	1,000,000	1,200,000 ^d

^aReported in *Electrical Merchandising*, January, 1944.

^bEstimated by the Bureau of Agricultural Economics, January, 1944.

^cEstimated by Textile, Clothing, and Leather Bureau, War Production Board, from industry reports to the Bureau of the Census.

^dData supplied by the Durable Goods and Products Division, from shipments reported by industry.

estimates. Some comparisons are shown in Table 3. Since national samples of about 5,000 interviews were used to represent 37,500,000 households, it is evident that the general agreement is remarkable, especially in view of the uncertainty which attaches to the independent estimates which are used for comparison. Not all estimates were as close as those shown in the table, but where there was wide disagreement there was usually an explanation. Purchases of alarm clocks, for example, were found to exceed by a wide margin the number which had been produced. This was accounted for, however, by the active secondhand market which developed when the shortage became acute.

The agreements such as those shown in the table can be reached only if two conditions are met: first, the sample must be properly selected; second, the interviewing must be faithfully conducted.

An illustration of the interviewing hazards which must be overcome if projections are to be made from a sample which has been carefully designed is found in the importance of interviewing people who are hard to find at home. Study of the data on interview results obtained from those found at home on the first call and those found at home only after repeated visits showed a considerable difference. [2] Not only did the replies of those reached only after repeated calls vary from those obtained on the first call, but the differences were not predictable. Later calls differed little for replies concerning electric irons, radios and refrigerators, but differed greatly for home ownership, washing machines, and sewing machines.

The experience of these surveys has shown their usefulness during wartime, and suggest that after the war similar studies might prove to be a useful service of government to business and the consumer.

REFERENCES

1. These estimates were reported in detail in "The Immediate Market for Appliances" and "The Market for Houses when Materials and Manpower are Available," released in October, 1944.
2. See the report by Ernest R. Hilgard and Stanley L. Payne. Those Not at Home: Riddle for Pollsters. *Public Opinion Quarterly*, 8 (1944), 254-261.

Chapter XXII

ATTITUDE SURVEYING IN THE DEPARTMENT OF AGRICULTURE

ALBERT A. CAMPBELL
Bureau of Agricultural Economics

INTRODUCTION

FOR many years the Department of Agriculture has felt it necessary to maintain a clear channel of communication between itself and the nation's farmers, not only so that it could transmit to farmers the many items of information with which they are concerned but also so that farmers could make their needs and attitudes known to the department.

The department has a great many different sources of information regarding the trends of thought among the rural population. The field representatives of its various divisions are stationed in every part of the country. These men are in constant contact with the people of their communities and are able to observe the manner in which the Department's programs are received by farmers. While the field people are usually concerned with some specific aspect of Department work, they serve as informal reporters of public opinion in their localities. These grass-roots observations have helped to keep the Department in contact with the people it serves.

In 1937, the first steps were taken toward the establishment of a more systematic procedure for measuring and reporting farmer attitudes. In that year a small group of interviewers were given the assignment of traveling around in the farming areas of the country and ascertaining as best they could the frame of mind of farm people in regard to certain Department programs which were being introduced at that time.

Despite many imperfections in procedure, the value of the information gathered in this way was so clearly established that the department felt justified in expanding its research on rural attitudes. In 1939 a new division, the Division of Program Surveys, was created in the Bureau of Agricultural

Economics for the purpose of carrying on surveys of the attitudes of farm people. Rensis Likert, a psychologist well known for his research in the field of attitude measurement, was brought in as its head.

In the ensuing years, the Division of Program Surveys has conducted a large number of interview surveys. These surveys are usually made at the request of some agency in the Department of Agriculture which requires information on public attitudes or information. Many studies have been done on contract for other government agencies. To a large degree, the Division of Program Surveys has become the central attitude-measurement agency of the government.

SUBJECT MATTER

Because it has been called upon to carry on studies for many different federal agencies, the subject matter handled by the division varies greatly from one study to the next. The following list of study titles indicates the diversity of the research with which the division is concerned and illustrates the kinds of problems for which interview research is conducted in federal administration:

- Midwestern Farmers' Appraisal of the War Food Program
- Victory Gardens and Home Preservation of Food
- Farmers' Views on Postwar Wheat Problems
- Attitudes toward the Accumulation and Expenditure of Liquid Assets
- Veterans' Readjustments to Civilian Life
- Consumer Acceptance of Dried Milk
- Land Market Participation by Factory Workers in Detroit
- Attitudes of Rural People toward Radio Service
- Public Response to the Seventh War Loan

Although the specific problems that the division is called upon to investigate vary widely and require various kinds of techniques, its surveys deal in general with three types of subject matter: attitudes, extent of information, and behavior.

Attitudes. Attitudinal surveying in the Division of Program Surveys goes far beyond the simple determination of the approval or disapproval by a representative sample of people of some government program. The agencies interested are rarely satisfied with only the distributions of the frequencies with which various attitudes occur. The emphasis in division work is on the explanation of the attitude. If, for example, the division undertakes to study the attitudes of dairy farmers toward the department's policy of paying a direct subsidy on milk to the farmer rather than permitting the retail price of milk to rise, it will explore the background of thinking from which the farmers' attitudes arise. Is opposition to the subsidy based on general resentment of government control, on impatience with the inconvenience involved, or on the feeling

that the subsidy restricts the farmer's legitimate profit? What reasons do farmers think the Department had for instituting such a program? Do farmers see any inflationary danger in an uncontrolled milk price? Do they think that the subsidy has influenced them to produce more or less milk than they would have produced without the subsidy? All of these questions are of great importance to the administrators of the department's milk program. The reactions of farmers to the program cannot be understood until these and other related questions are answered.

Answers to "why" questions are especially important in cases where the public's reaction to a program has been un-cooperative or hostile. Many federal programs are based on the voluntary cooperation of the people whom they affect; if this is not forthcoming, it is imperative that the officials responsible for the program should know why. Division surveys have explored a number of such problems: Why did farmers in the Corn Belt not respond to the War Food Administration's goal for hog production? Why did fuel oil users fail to apportion their rations properly over the winter? Why have farmers in certain areas taken to burning down stands of timber in national forests? Why did potential War Bond buyers fail to respond to some aspect of the promotional campaigns? Effective administration requires that an explanation for these failures be obtained.

In some cases the division is asked to assess attitudes toward programs that have been proposed but have not yet been put into operation. Studies of this type have been of especial importance in relation to new programs contemplated for after the war. What, for example, are the attitudes of cotton farmers toward the control of possible cotton surpluses after the war? Do cotton farmers feel that they would be better off if the government withdrew entirely from the cotton market and permitted cotton to find its own price and production levels? If they object to government participation in the cotton market, on what grounds do they base their opinions? If they feel that the government should continue to exercise some control over cotton, what kind of action would they like to see? Do they favor acreage restriction, price supports, direct supplementary payments? How would they feel about a program of government purchase and resale of surplus cotton? What expectations are these preferences based on? Do cotton farmers expect a surplus or do they think that the domestic and world markets will absorb the entire production? Do they think cotton prices will fall if no restrictions are placed on production?

An investigation of these questions and others like them provides the Department of Agriculture with a better understanding of cotton farmers' thinking regarding these important postwar issues and assists in the selection of a program that will meet the farmers' needs and preferences as closely

as other considerations will permit. Many factors aside from public opinion must be considered when a course of action is being planned; nevertheless, few administrators care to disregard public attitudes when it is possible to adapt their programs to them.

Extent of Information. While attitude research is the principal concern of the division, frequently its surveys contain questions designed to ascertain how well people are informed about specific aspects of federal programs. Such data are often essential for the full understanding of public reaction to the programs. It is impossible to know, for example, to what extent the nation's farmers approve of the so-called parity price system unless it is also known how well they understand it. If "parity" is known only as a poorly defined device for keeping up the prices of farm products, the approval of it which farmers may express is very much less meaningful than if they approve it with a full knowledge of the manner in which the parity system attempts to equate farm and city living standards.

Data regarding public information are often desirable to assess the effectiveness of campaigns which have been carried on by federal departments to inform the public regarding their programs. If the Department of Agriculture is mailing out to farmers periodic information on the current food situation, it needs to know whether these releases are being read and understood. The department needs to know whether the radio programs it sponsors are interesting to farmers and whether farmers are profiting from them. Other federal agencies have similar problems. The Treasury Department must find out how well its War Bond drives have reached the public, how many people knew their local quotas, how widely known the different selling slogans were, how well people understand the purposes of bonds, and the answers to other similar questions. The Office of Price Administration must know whether its directives are being understood by the wholesalers and retailers they affect. The State Department needs to know how successful it has been in explaining the difficult problems of international organization to the public.

Behavior. Most division surveys, although essentially attitudinal in character, are partly directed toward the accumulation of purely behavioral data. How many people have planted home gardens? How many people have opened checking accounts in the last three years? How many people bought War Bonds in the last campaign? How many people share their cars in driving to work? How many people have bought dehydrated food during the war period and what foods did they buy? Many questions of this type have been asked in division surveys, almost always with accompanying questions designed to ascertain why the respondents had or had not taken these actions.

Typically, the administrative needs that give rise to the surveys made by the division require information of all these three types. Although the divi-

sion's major interest is attitude measurement, none of its surveys is purely attitudinal. Information about what people know and what they have done is almost always essential for a full understanding of their reaction to the various federal programs.

PROCEDURES

The procedures followed in making these surveys vary according to the character of the study to be done. Since division surveys range from investigations of attitudes toward the proper role of government in postwar agriculture to inquiries about the acceptability of dehydrated foods, it is clear that adaptations to specific subject matter are necessary.

The planning of a survey begins with a careful definition of the objectives to be covered. This goes beyond a general statement of the purposes of the study; its final stage is a detailed listing of the specific data to be obtained. If the study in question is being done at the request of some agency outside the division, the planning is done in co-operation with representatives of this agency. Their special information regarding the program to be investigated is essential at this stage of the development of a survey.

After the specifications of data have been completed, a questionnaire is written, designed to provide information on each of the specified items. The wording and sequence of the questions to be used are selected with great care to avoid possible misunderstandings or undesirable effects resulting from emotionally toned words or phrases. Before this questionnaire is actually put into production, it is given a pretest in the field. The preliminary trial has two general purposes: first, it provides a test of the clarity of the questions and the correctness of the interpretation put upon them by the respondents, and second, it makes possible the discovery of new aspects of the problem being studied that were not anticipated in the planning stage. Both these functions are highly important as protection against incompleteness and confusion in the final results of the survey.

The Division of Program Surveys is unique among agencies doing interview surveying in its highly developed use of the so-called "open question." Most of the questions used in the division's interview schedules are phrased in a conversational tone, and no answers are specified for the respondent's choice. The respondent answers the question in his own words as he would in an ordinary conversation, and his answer is taken down, virtually verbatim, by the interviewer. Through the use of questions of this type the division feels that it avoids the dangers implicit in suggesting plausible answers to the respondent and also that it obtains a much fuller understanding of the respondent's attitudes than can be achieved through use of questions of the checklist type. Typically, an interview will be a half hour to an hour's conversation in

which the interviewer asks a specific series of carefully worded questions and the respondent replies in his own way, the answers being recorded as faithfully as possible by the interviewer. Division interviewers are specifically trained in the art of "nondirective" probing by which they draw a full answer from their respondents without suggesting possible answers or adding new questions not specified in the interview schedule. These interviewers are on the alert for significant attitudinal comments their respondents may volunteer on government programs that are not the direct subject of the interview. Thus, developing public dissatisfactions can be detected in their early stages.

Not all of the surveys conducted by the division require the exclusive use of the "open question." In many instances in which the possible answers to a question are limited and clearly understood checklist questions are used. The major determinant of the type of question used is the character of the subject matter to be covered. Generally speaking, the "open question" approach is used when the objectives of the survey bear on complex attitudinal problems or areas of thought in which public information is limited.

While the division utilizes a variety of sampling procedures, depending on the nature of the universe being considered, its basic procedure is area sampling. In following this technique, the interviewers do not select their own respondents in order to fill specified quotas of individuals of different types. On the contrary, they receive specific instructions regarding the households they are to visit. These households have been previously selected by application of stratified-random sampling methods so as to give a good scatter over the population being studied.¹

When conducting surveys in rural areas, use is made of the Master Sample of Agriculture, which has been developed by the Bureau of Agricultural Economics and the Bureau of the Census. The Master Sample consists of approximately 70,000 small geographical areas, selected from every county of the country. Subsamples of these areas may be drawn giving either national or regional coverage.

In nearly all division surveys the interviewers are required to make at least two attempts to interview the originally designated respondents before they make substitutions. It has become apparent that people who stay at home a good deal of the time may differ both in attitudes and in activities from people who are hard to find at home, even though they may be similar in age, education, and other factors sometimes used in determining sampling quotas. For example, the division has found that respondents who are at home the first time the interviewer calls are considerably more apt to have home gardens

¹ For a full explanation of these sampling procedures, see King, A. J., and Jessen, R. J. The Master Sample of Agriculture. *J. Amer. Stat. Assoc.*, 40 (1945), 38-56. See also Chap. 20 in this volume.

than are people who are found at home on the second call; these in turn are more apt to have home gardens than are people who are interviewed only after three or more calls. In other words, the number of home gardens in the country would be considerably overestimated if the estimate were based on a survey in which the interviewers were allowed to call from door to door until they found someone at home. Because of differences of this sort between people who stay at home and those who do not, the division goes to considerable effort to locate its sample as originally defined and insists that when substitutions must be made they be designated in the same systematic way as the respondents for whom they are made.

The size of the sample used in any particular survey is determined by several different factors. A heterogeneous population requires a larger sample than one which has relatively little range of variation. Samples that are not going to be broken down into subsamples in the analysis do not need to be as large as those for which such fractionation is contemplated. Surveys done for the purpose of achieving a rough measure of the distribution of attitudes do not require as large samples as do studies that are expected to yield estimates of— for example—the total monthly income of some part of the population.

Because of these and other considerations, the size of the samples used in division work varies considerably. For nationwide surveys of purely attitudinal character samples of approximately 2,500 are commonly used. This permits comparisons of the attitudes of men and women, poor and well-to-do, and other large groupings of this sort. It does not, of course, permit statements regarding individual states or other small subdivisions of the total sample. Comparable studies of restricted populations, such as dairy farmers in the Great Lakes area, would not require more than four or five hundred respondents. On the other hand, a survey that proposed to estimate the total amount of money held in liquid assets in a single county might require a sample of 400 or more respondents, with extra loading in the upper income brackets to assure an adequate representation of that small but highly important group. As these examples indicate, there is no rule of thumb regarding sample size; it is determined in each case by the specifications of the individual study.

APPLICATIONS

The applications of the findings of the division's surveys are both specific and general. In some cases survey results are transmitted directly into administrative action. The discovery, for example, that a department directive limiting the use of certain forest lands contributed to an outbreak of man-made forest fires in a local region brought about an alteration of policy in that area that was later expanded into a revision of policy throughout the country. The

finding that a particular promotional release issued by one of the federal agencies had failed to reach the audience to which it was directed resulted in the abandonment of that program and the substitution of more successful procedures. The fact that dairy farmers were found to be reluctant at the beginning of the war period to increase milk production because of their uncertainty regarding future milk prices brought about a refocussing of the department's campaign to increase dairy production with a greater emphasis on the stability of its price support program for milk.

While most of its surveys bear on specific administrative problems of this type, some of the work of the division has a more general character. A number of studies have been done, for example, to ascertain the role the nation's farmers would like to see the Department of Agriculture play after the war. These studies have not been directed toward the immediate solution of some specific problem of administration; they are intended rather to aid the department in its long-range planning of postwar policy. They are meant to portray to the policy makers of the department those basic attitudes of farmers regarding postwar issues which must be considered in determining postwar action. Another series of surveys has been conducted for the purpose of following the development of farmers' attitudes toward land prices and participation in the land market. These studies were not done because of an immediate need for executive or legislative action. Their objective is to provide a barometer of the intensity of "boom psychology" among farmers and to obtain the opinions of farmers about the relative acceptability of various programs of control of land buying which might be undertaken. If land inflation should reach a point at which the department felt it should take or recommend action, it would have at hand this significant information to aid in formation of policy.

The relative value of these two types of surveys—those that bear on immediate problems and those that anticipate problems—is difficult to assess. The applications of the findings of surveys done to meet immediate administrative needs are usually clearly demonstrable and their value can sometimes be computed in dollars and cents. No precise accounting of this sort can be made for the more general surveys that are directed toward broader objectives and less direct application. However, despite the difficulty of establishing the monetary value of surveys of this character, their importance in helping complete the total background of information on the basis of which the Department takes action is seldom questioned.

AN EXAMPLE

In connection with planning the postwar activity of the Department of Agriculture, the division undertook (in early 1944) a survey of the thinking of

farmers in different parts of the country. This concerned some of the important problems they would face as agriculture returned to peacetime production. The purpose of the study was twofold: to provide the basis for an information program designed to increase the understanding of farmers of the economic issues involved in the reconversion to peace, and to give department administrators a statement of the part farmers wanted to see the department play in postwar agriculture.

In order to meet these two objectives it was felt that the survey should cover the following groups of items:

1. Farmers' expectations regarding the level of farm prices after the war, their attitudes toward wartime price controls and the desirability of continuing such controls after the war, their expectations of the probable consequences of removing price and inflation controls when the war is over.
2. Farmers' expectations about general economic conditions after the war, their suggestions about ways of avoiding widespread unemployment at that time, their beliefs about the capacity of private business alone to accomplish the transition from war to full peace-time employment.
3. Farmers' attitudes toward the desirability of using production adjustment programs to help shift agriculture back into peacetime production.
4. Farmers' attitudes toward public works programs as a means of relieving unemployment, their estimates of the particular forms of public improvement which would be most desirable from the viewpoint of the community development.
5. Farmers' expectations about market conditions of farm products after the war, their attitudes regarding active United States participation in world trade, their attitudes toward various surplus disposal programs, their ideas about the proper course of government action in the event of a postwar depression.

The questionnaire devised to draw out these attitudes, expectations, and beliefs consisted of a schedule of some 35 open questions. An additional list of census-type items, such as, sex, race, age, education, income, and the like, was appended to provide background information which would be useful in the analysis of the attitudinal data. The following excerpts from actual interviews illustrate the type of open questions used and the kind of response which they elicited:

What do you think would happen to prices (of farm products) in general if price control programs were done away with right after the war?

If done away with immediately, it would have a terrible effect on prices. It would run prices down and denormalize people and run them out of business. They ought to be taken off gradually, worked off gradually like they were worked on gradually. In case farm prices drop sharply after the war, should the Government step in and do something about it?

Yes, we could not stand supply and demand unless other things drop with our stuff. We farmers just live from one crop to another and something like that would ruin us. The Government ought to hold up stuff so that our stuff would not drop too fast in proportion to other stuff. I am in favor of the Government letting more responsibility go to the farmer, put him on his own responsibility, but that will have to be done gradually because we have gone into all this gradually.

What sort of things should the government do?

The only thing that they could do would be to carry on the subsidy payments for a while. I am not much in favor of that but that will be necessary while we change back. It got to where we could not plant over 25 per cent of our land in cotton, and before that started some planted 50 to 75 per cent of their land in cotton and did not make any feed or anything to live on. When we cut down under that program we began using fertilizer on the cotton, and a good farmer could make as much as he did before on more acres, and he raised about as much feed and stuff as he needed. That was a good thing.

Should the Government help farmers plan what to produce after the war, or should farmers be left to produce what they please?

No, farmers ought to be free to plan for themselves, I think.

A somewhat different point of view is expressed by another farmer in answering the following questions:

Do you think farmers will make less money if people in the cities don't have jobs? Why?

Sure, because if people in the cities don't eat meat and things, the farmer can't sell them. They spend the money they can for food, and if they don't have it we don't have a market for it.

Do you believe that businessmen can provide jobs for everybody after the war without help from the Government?

I doubt if they can do it even with the help of the Government. When everybody was out of a job before, the government put them to work on the WPA and that helped a little. But there were still some on relief and I think it will be the same again. They won't be able to put everybody to work.

In case a good many workers are out of jobs after the war, do you think the Government should provide jobs for them?

If it's anyways possible. If they are willing to work and are not dead-beats on the job. What do you think of the way the Government put people to work during the depression?

They didn't earn what they got but it was better than the soup line. It made them appreciate what they got more. I think they were paid more than was necessary in that depression time and for the work they did. But it was a lot of help to the poor people and they wouldn't have had anything if it hadn't been done. Some didn't deserve it but there are always those in everything, if there is some way of getting around work. But taking it all the way around, it was all right.

The survey was restricted to farm operators in three areas: the Corn Belt, the cotton region of the Southeast, and the Central Valley of California. These three areas represented widely different kinds of farming problems. Samples of approximately 300 respondents were interviewed in the Corn Belt and in the Southeast and a somewhat smaller number in California. Each sample was selected to represent all farm operators in its particular area.

The findings of this survey followed the outline of objectives drawn up in the planning of the study. A partial statement of the summary of the report follows:

Farmers in the Corn Belt, the Southeast, and the Central Valley of California commonly expect that there will be widespread unemployment after the war, and they believe that the Government ought to provide jobs when business is unable to do so. They are not opposed to employing the jobless on public works, but they do believe that some of the defects of the pre-war work programs should be corrected.

Farmers in the Corn Belt and in California are fairly optimistic about income prospects for the first two years after the end of the war in Europe, because they think foreign demand will sustain their prices; but they expect a depression eventually. Southeastern farmers more often expect depression to set in immediately after the war.

Farmers in the three areas do not want the Government to discontinue its planning of the farm economy as soon as the war is over. The Southeast particularly wishes continued price and production planning in agriculture.

Because of the unusual interest of this study, a direct report of the findings was presented to the Secretary of Agriculture and his staff, and their implications for departmental policy were discussed. The report was then circulated in the Interbureau Committee on Post-war Programs for its consideration and for integration into the general recommendations for policy with which this group was concerned. Subsequently hundreds of copies of the report were distributed within the various governmental agencies, to the many field offices of the department, to numerous agricultural colleges, research agencies, and other interested organizations.

In order to exploit more fully the educational value of the study, the Office of Information of the department rewrote the original report in the form of an illustrated pamphlet, of which thousands of copies were issued through the state offices of the Extension Service. Special short releases summarizing the findings were issued for the use of rural newspapers throughout the country. A discussion of the survey was recorded in Washington, and reproductions were made available to the numerous rural radio stations which feature programs for farmers. The natural interest of the subject matter of the study has given the report unusually wide circulation.

CONCLUSION

The Division of Program Surveys is called upon to conduct surveys of the kinds mentioned here because the Department of Agriculture and other federal agencies feel that efficient administration demands information of this kind. Just as business organizations find that market research is necessary to their effective operation, government offices also require the services of an agency which can systematically report the state and trend of public attitudes and information. The usefulness of such reporting in assessing old programs and planning new ones is at least as great for government officials as it is for business men. Both are responsive to the public wish; both feel the need to be informed about it.

PART THREE

SURVEYS REPORTED TO THE PUBLIC

Chapter XXIII

TRENDS IN PUBLIC OPINION RESEARCH

WROE ALDERSON

Wroe Alderson, Simon, and Sessions.

THE origins of the public opinion polls go back to the early twentieth century. At first the polls were conducted by newspapers and confined to their own immediate areas. For a long time they were strictly journalistic in conception and objective. A newspaper editor undertook a poll of opinion concerning candidates or issues to provide editorial material for his newspaper with a high degree of reader interest. In many instances he was consciously or unconsciously motivated by a further desire to influence the result of the election toward the side which he favored. Rival newspapers polling their own communities often got opposing results, each claiming meanwhile to be using fair and impartial methods.

Journalistic motivations of public opinion polls continued to be the primary consideration when the idea was taken up by national publications. *The Farm Journal* was the pioneer of public opinion polls on a national scale, having engaged in polling as early as 1912.

Therafter the best known poll operating over a period of years was that conducted by the *Literary Digest*. The final *Digest* poll was the one in 1936 which was wrong to such a disastrous extent as to serve as a good springboard for the new organizations using more scientific polling techniques.

The three poll-taking organizations which got under way with the 1936 national election, and have continued to be outstanding since that time, were founded by three research men who had all had experience in the use of sampling methods. Two had been interested primarily in reactions to consumer products, and the other in newspaper reading habits. These men are Elmo Roper, Archibald Crossley, and George Gallup.

There are some sharp differences in points of view both as to what the polls are designed to accomplish and as to the best means of achieving these

results. They differ on the most fundamental question of all: whether the polls should concentrate on measuring public attitudes or whether they should accept responsibility for predicting election results. While these objectives are necessarily related, emphasis on one or the other makes a great deal of difference in the polling techniques and statistical procedures employed.

Two of the organizations are clearly identified with these opposing views concerning the functions of the polls. Roper is concerned solely with measuring national sentiment on an issue or a candidate. Thus, in a national presidential election, he considers his sample an indivisible national cross section and makes no attempt to predict the electoral votes from the analysis of returns state by state.

Crossley has stood for the view that the purpose of the polls is prediction and that they should use whatever sampling, interviewing, or statistical techniques are necessary in order to forecast the election results. Thus, the Crossley organization typically makes all or most of its interviews in critical or doubtful states, generally ignoring, for example, the states of the "solid South" where the result is pretty well settled in advance.

The viewpoint of Gallup's organization lies somewhere between these two positions. The difficulty this organization has had in choosing between the viewpoint of measurement vs. prediction undoubtedly arises partly from the fact that the newspaper reader for whom the polls are made does not make a clear distinction between these objectives. He assumes that percentages reported by a polling organization represent unadjusted percentages showing the portions of the electorate adhering to specific attitudes. He also assumes that these attitude measurements represent the poll taker's implied prediction of the outcome. Because of its greater concern with this problem of presentation, the viewpoint of the Gallup organization may be described as journalistic.

A brief description of each of these organizations will now be presented, emphasizing in turn the problems of scientific measurement, scientific prediction, and the journalistic presentation of results. This discussion of the three leading polling organizations will be followed by a consideration of two organizations of quite different character devoted primarily to the improvement of polling techniques and the general advancement of public opinion research. These latter organizations are the Office of Public Opinion Research at Princeton, N. J., and the National Opinion Research Center at Denver, Colo.

THE ROPER POLL AND THE MEASUREMENT OF ATTITUDES

Roper's first public opinion poll was taken in the spring of 1935 and the results were published in the May issue of *Fortune*. Thus, the *Fortune* poll

may claim a few months' seniority among the three leading public opinion polls, although national attention first became focused upon them with the election campaign of 1936.

The principles of scientific sampling adopted by Roper and the other poll takers were already well established and had been used successfully in market research for some quarter of a century. The fundamental principle is simply that a sample must be representative and that the question of size of the sample is a secondary consideration. A perfectly representative sample is one which is a perfect model of the total population to be sampled, paralleling the larger group in all its principal dimensions and proportions. The national sample taken by Roper in 1936 was only 3,500, although the number has since been increased to 5,200. Since this was the smallest sample taken by any of the three leading pollsters, it represented the most confidence in the validity of representative sampling.

There are two general methods of drawing a representative sample from a population. One is to make a purely random selection out of all the members of the larger group. The other is to build up the sample as a working model of the larger group on the principle of stratification. All of the poll takers chose the method of stratification. The reasons for the choice are obvious. In the first place, no consolidated list of all persons eligible to vote in the United States was available. In the second place, after a sample had been drawn from such a complete list, the cost of reaching the individuals through personal interviews would have been prohibitive.

The results obtained by the new polling organizations with their small national samples was a clear triumph for the principle of representative sampling, but it did not necessarily mean that Roper and the other poll takers had employed any difficult or exceptionally refined techniques.

The most elementary approach to representative sampling was enough to improve on what the *Literary Digest* had been doing. The first weakness of its procedure was in drawing up its lists from such sources as automobile registrations and telephone books. Thus the sample from the very beginning was unbalanced from the standpoint of economic status. The fact that ten million cards were mailed out to people on these lists made it an impressive effort for most readers of the magazine. Unfortunately ten million was just as poor a sample as ten thousand so long as the procedure contained this fatal bias from the start.

The second weakness was in relying on mailed ballots rather than personal interviews. The mailed ballot gives undue emphasis to the element of self-selection arising from intensity of feeling or other factors which might cause one respondent to return his ballot while another failed to do so. Because of this factor it could easily happen that an embattled minority representing one-

third of the electorate would account for two-thirds of the ballots returned. This factor of variation in responsiveness is still one of the unsolved problems of polling techniques. Its influence, however, was greatly narrowed in one great jump by substituting personal interviews for mailed ballots.

Roper has had an edge on the other pollsters in measuring the size of the popular majority in national elections. This is not due to either size of sample or principles of sampling, since all have employed the same general pattern in building stratified samples. Roper has possibly used somewhat greater care in training and supervising field interviewers. The interviewer must be both honest and skilled in selecting individuals in accordance with a sampling pattern, or no plan of stratification will work. The chief difficulty in the interviewer's selection of respondents in accordance with quotas has to do with the economic status of the family. Roper appears to go further than the other organizations in setting up elaborate descriptions of families in its economic classifications and in training interviewers to make accurate selections on this basis. The difference in this respect is apparently not decisive, since Gallup and Crossley have obtained closely similar percentages in each election.

The most characteristic feature of the Roper technique is the use of the attitude scale in interviewing. Under this technique the respondent is presented with four or more statements and asked to indicate *which most nearly represents his attitude concerning a candidate or an issue*. One objective of this procedure is to cut down the number who claim to have no opinion or who refuse to co-operate altogether. Roper believes that some people who are reluctant to express themselves positively on one side or the other will more readily indicate an answer that shows they are leaning in one direction or the other. Alternative methods of interviewing will be discussed more fully later.

THE CROSSLEY POLL AND ELECTION FORECASTS

The Crossley organization took its first national poll in 1936 after having established itself as a pioneer in the measurement of audience reaction to radio broadcasts. In his radio surveys and in other marketing studies, Crossley has generally held that exact measurement is secondary to prediction or to the decisions which are to be taken on the basis of research results. In his public opinion polls, he considers the ultimate objective to be prediction of the election outcome. Thus, in 1940 the Crossley poll covered only 19 states which were considered to be the pivotal and were expected to decide the election. His forecast of electoral votes in that year deviated from the actual election returns by only 1.8 percentage points.

In 1944 the Crossley organization was the only one to attempt to estimate the vote of the armed services. His estimates of the serviceman's vote made

before the election indicated that it would add 0.8 percentage points on the Roosevelt side, varying, of course, from state to state. This is in line with Crossley's general conception that he must attempt to measure the actual turnout state by state since the votes in the electoral college are determined for each state as a whole.

Because of his emphasis on prediction, Crossley is interested in possible discrepancies between the preference indicated in a poll and the way people will actually vote. He was the first to experiment with the secret ballot. Crossley abandoned the use of the secret ballot in 1944 and employed four separate tests of the degree of voting interest. All these tests fortified his conclusions in his final and definite forecast that Roosevelt would be elected. In the first test people were asked to name the specific reasons that had led them to their choice of a candidate. The character of the reasons represented were held to indicate a more positive and deep-seated pro-Roosevelt sentiment than pro-Dewey sentiment.

A second method was to ask people in September and again in October how sure they were that they would not change their minds before election. The answers showed a greater increase in certainty among those who were going to vote for Roosevelt as against those who were going to vote for Dewey. A third check was to ask respondents how long ago they had made up their minds and to analyze the results by decision periods. Among those whose decisions had stood for the longest period of time, 52 per cent were on the Roosevelt side.

The fourth test was a special attempt to measure the degree of voting interest, or, stated in reverse, the degree of apathy which might keep either Roosevelt supporters or Dewey supporters away from the polls. This test began with a study of qualified voters, who were asked about their period of residence at their present locations and comparing the answers with voting regulations. By this and other tests, those who would not be eligible to vote were weeded out. Interviewers then went back to the eligible group and asked them a series of questions designed to reveal the likelihood of their actually voting. This test showed a somewhat greater determination to vote among qualified voters who said that they favored Roosevelt. This kind of precaution appears essential for an organization which undertakes to predict the election results. These methods no doubt accounted for the relatively high degree of accuracy in Crossley's state-by-state predictions.

The Crossley organization was closer to the actual result than any other major forecast of the election. In his final prediction Crossley underestimated Roosevelt electors by 78; Gallup, by 140. Roper made no prediction on the electoral vote.

Other attempts to forecast the election outcome did not rely primarily on public opinion polls. *Newsweek* magazine used a panel of 118 political writers.

The consensus of this group forecast a Roosevelt victory, but with a greater error in the state score than either Crossley or Roper. Another forecaster, using recent election trends state by state, forecast the outcome in 1944, but not so accurately as Crossley. Thus, the measurement of public opinion through the polling technique seems to be established as the most dependable foundation for successful prediction. There may be some comfort in this fact with respect to our somewhat faulty electoral machinery. Apparently the winning candidate is usually the one that the majority of citizens favor despite the failure of many to vote and despite the paradoxical results that theoretically can emerge from the electoral college system.

THE GALLUP POLL AND ITS JOURNALISTIC PRESENTATION

Gallup set up his polling organization in the latter part of 1935 and achieved national fame in the presidential campaign of 1936. The Gallup poll in that year used a sample of 40,000 and forecast the popular vote with an error of 6 per cent as compared with 19 per cent for the *Literary Digest* and 1 per cent for Roper. The Gallup poll was neither the first of the modern polls nor the one to make the most striking demonstration of the use of small samples. Since its founding it has not been the most accurate in measuring the popular vote nor the most successful in predicting the electoral results. What, then, has made the name of Gallup the universally accepted symbol for the public opinion polls? The answer lies to a large extent in the arts of journalistic presentation of the results. The problem of presentation will be emphasized in discussing the Gallup poll, but without any intention of minimizing the continuing contribution of Gallup and his organization to the art and science of public opinion research.

The arts of public presentation as practiced by Gallup start with a good journalistic sense of what will interest the mass of people and what they can understand. Gallup's statements have usually been models of simple and lucid expression. Simplicity in technical matters must always be achieved at some cost in the suppression of significant detail. While there may be other reasons for preferring questions that require simple yes or no answers, it is certainly easier to present the results to a newspaper audience than to interpret the outcome of the more elaborate multiple-choice and open types of questions.

Another aspect of this journalistic viewpoint, which has already been mentioned, is the fact that Gallup does not, like Roper or Crossley, take a clear-cut position in the direction of either measurement or forecasting. In general approach the Gallup organization appears to emphasize the aim of measuring public opinion. At the same time, there is a realization that the public expects accurate forecasting and that the continued prestige of the polls rests largely

on their success in this direction. As a scientist, Gallup knows that his poll could sometime be wrong under the laws of statistical probability. As a journalist he feels that he must always be right in the minds of his readers. This journalistic requirement has led to two unfortunate practices. One is an extreme form of hedging in final forecasts, the other is making specious claims of accuracy following elections.

The final forecast of the Gallup poll is usually so hedged that all the preceding field work and analysis seem futile. The accuracy claims are overstated and not warranted by the facts.

For instance, following the 1944 election, the American Institute of Public Opinion claimed it had forecast 446 Roosevelt electors, as against 432 elected. Far from indicating such a landslide, Gallup's prediction led the reader to believe a Dewey victory a distinct possibility. The total of 446 votes includes not only the states Gallup had predicted as definite and leaning toward Roosevelt, but also 154 votes defined as leaning toward Dewey. The only way one could obtain the prediction Gallup's post-mortem claimed was to disregard the interpretive text of the last poll, and to shift the votes leaning toward Dewey to Roosevelt on the ground that the Gallup poll always had underestimated the Democratic vote!

On the other hand, Gallup has a keen understanding of the problem of public presentation and has accomplished much in establishing general acceptance and maintaining the prestige of the public opinion polls as a whole. His foresight and generosity with respect to such nonprofit organizations as the Office of Public Opinion Research is to be commended.

Recently Gallup has announced a radical change in sampling techniques which he says will increase the validity of the poll results. Actually what he is talking about is dropping the principle of stratification and going over to a form of random sampling—the "area sample" developed by the Census Bureau and discussed in Chapter XX. There is some question of whether the time is ripe for the complete adoption of the census method by the poll takers, since they customarily use much smaller samples than are used by the Census Bureau.

PUBLIC OPINION RESEARCH AND SOCIAL PSYCHOLOGY

There are two permanent agencies exercising functions of criticism and review of public opinion polls, both operating under foundation grants. The Office of Public Opinion Research began in 1940 and the National Opinion Research Center in 1941. Both are concerned with advancing the science of public opinion research. There is a good deal of overlapping in their activities. At the same time, there is a real difference in their major emphasis, with the

Office of Public Opinion Research representing primarily the viewpoint of social psychology and the National Opinion Research Center that of political science. The first organization is concerned with the formation and structure of public opinion, while the second appears more vitally interested in the relation between public opinion and the procedures by which actual political decisions are taken. These two aspects of opinion research as a branch of social science will be emphasized in describing the work of these two organizations.

The Office of Public Opinion Research at Princeton was established through the initiative of Gallup and is headed by Hadley Cantril, professor of social psychology in Princeton University. Cantril has worked intensively on the psychological problems of question wording, on sources of hidden bias in opinion research, and on psychological and social factors influencing the formation of opinion.

He has made a great variety of experiments with the three principal types of questions which have been used by the polling organizations. These three types are known as the free or open question; the yes and no, or dichotomous, question; and the multiple-choice, or cafeteria, question. The open question is one in which respondents are given an opportunity to say what they think on an issue, putting it in their own words rather than merely choosing between suggested answers. The dichotomous question is one that asks the respondent to register a definite preference on a stated issue or between two principal candidates. One example is the poll in which people were asked whether they expected to vote for the Democratic or Republican candidates. The multiple-choice question is one which presents several possible answers in an attempt to obtain a graded response. One type of multiple-choice question sets forth a series of statements and asks the respondent to indicate which one most nearly indicates his attitude. The application of this method is most obvious in measuring opinion on issues, since individuals may be for or against a proposal with reservations rather than taking an unqualified stand on one side or the other. This method has also been used in polls on candidates, giving the respondent a chance to indicate not only his choice but the intensity of his feeling on the side of his choice.

Aside from the type of question, Cantril has been concerned with the influence of words which have acquired emotional significance with respect to a public issue, and hence should be avoided in polling questions. The inclusion of a word that suggests disapproval of a given position will cut down the percentage expressing that opinion as their own. Similarly, a word suggesting approval will attract the votes of respondents who would otherwise express no opinion or who are weak in their convictions.

Cantril reaches several general conclusions on the general problem of question wording. The first is that people whose opinions are strongly held

and have been worked out in detail are not readily influenced by the form of question. A large percentage of respondents do not fall in this classification, so that the objective wording of questions is an essential part of polling techniques. As to the three general types of questions, it is not correct to say that one is superior to the others, since there are special spheres of usefulness for each type. The open question has great value in exploration. Its use is indicated when a public issue has not yet been sharply drawn or when the poll takers are not sure of the way in which the issue has been formulated in the public mind. Open questions have made great contributions in market research, frequently enabling the researcher to discover aspects of demand which he would not have been able to foresee and to cover in the form of precise questions. It has similar advantages in some matters of public opinion.

The dichotomous question has special application in just the reverse type of situation, where the issue is already sharply drawn as between the two major candidates in a presidential election. It has the great advantage of simplicity both in obtaining information from the respondent and in tabulating and presenting the results. The method has been criticized on the ground that it may result in too large a percentage who refuse to answer or who state that they have no opinion.

The multiple-choice type of question was devised largely to cut down the size of these two categories. Roper has been one of the principal proponents of this type of question and, judged by the accuracy of his results, it has worked well for him. Others feel that in less skillful hands the cafeteria question has serious drawbacks. First is the logical difficulty sometimes encountered in establishing clear-cut alternatives which are mutually exclusive and yet cover all the possibilities. Another is the danger of confusing or irritating the respondent so that the results, even though easily tabulated, may not be reliable.

Other variations in interviewing method include the so-called split ballot, in which the same question is asked in two different ways with separate groups of respondents. This method appears to be nothing more than telescoping a testing procedure that might have taken place over a time interval. Whatever method is used, pretesting is generally accepted and regularly employed by all the principal polling organizations. Much work remains to be done in improving the several methods of interviewing as tools of research and using them skillfully in appropriate polling situations.

One of the most useful lines of research pursued by the Office of Public Opinion Research is its endeavor to evaluate the various sources of hidden bias which may distort measurements of public opinion. The great and continuing hazard of all forms of social research is that they depend upon the co-operation of the people who are being studied. In public opinion research co-operation is sought in answering questions and in making true statements of opinions.

The division of those who are approached in a public opinion poll into cooperators and non-cooperators presents an intriguing puzzle at the very start. It is impossible to determine precisely how representative is the opinion of those who cooperate, since forming this judgment necessitates knowing the opinion of the non-cooperators who refuse to give it. Cantril dismisses this difficulty rather too easily when he says that a series of tests show only 14 per cent refusing cooperation. There is no assurance against the possibility of the 14 per cent mostly running in one direction. This degree of non-cooperation could easily upset the results where the measured difference was only 3 or 4 per cent as it usually is in national presidential elections. There is the further danger that rebuffs may react on interviewers so as to make them avoid the type of neighborhoods or the type of people from whom they come to expect refusals. In a given instance it might turn out that the percentage of refusals was much higher than the stated figure of 14 per cent. If it runs too high it may indicate that something is wrong with either the interviewer or the question. A precaution, particularly in the testing of a new question, is to have interviewers report the number who refuse to co-operate and possibly to state whether refusal was connected in any way with the nature of the inquiry, or whether it was a more casual reason such as being too busy to talk. More work needs to be done in evaluating the possible influence of non-cooperation. A possible method of study which is not reported by Cantril would be that of compiling a list of people whose opinions were known through other channels and then having interviewers attempt to secure their co-operation in stating their opinions.

Cantril has reached more definite conclusions in studying distortions that arise in relationship between interviewers and co-operating respondents. Experiment with secret ballots as compared with oral interviews have shown that respondents are not always frank in stating their opinions. An unpopular opinion or one that reflects in any way upon the prestige of the respondent often gets a higher rating in the secret ballot than in oral replies.

Cantril divides the problem of distortion into the two phases of bias and rapport. He uses the word "bias" to designate distortion arising from strongly held opinions on the part of the interviewer. The influence of the interviewer's opinions does not appear to be reduced correspondingly by the amount of experience which the interviewer has had. The principal remedy so far available seems to be that of maintaining a balance of viewpoints within the staff of interviewers. Thus Gallup uses approximately the same number of Republican and Democratic interviewers. Rapport is not a matter of the interviewer's opinions but of bonds of sympathy and feelings of like-mindedness between the interviewer and the respondent. Negro interviewers obtain what appear to be more plausible answers than do white interviewers in deal-

ing with Negro respondents on questions of race relations. Similar distortions may occur where interviewer and respondent are of different socioeconomic status.

One interesting result obtained by Cantril is the apparent increase in the influence of the interviewer's opinion as the size of town decreases. The practical conclusions that he draws from these results are questionable, involving some inconsistency in his conceptions of bias and rapport. He suggests sending interviewers from near-by cities to talk to people in rural areas to eliminate the likelihood that the interviewer through personal acquaintance in the community would approach chiefly people who held the same opinion as himself. As a matter of fact, a great proportion of interviews in rural areas are made by interviewers working out of larger places. Thus the observed distortion may result from lack of rapport and an unwillingness to speak frankly to a stranger or a tendency to say whatever seems acceptable to the interviewer in order to terminate the discussion. Bias and rapport along with the problem of non-cooperation need further study in order to achieve greater precision in polling procedures.

A third field in which the Office of Public Opinion Research has made an outstanding contribution is in its findings concerning the major determinants of public opinion. The factors considered include age, sex, race, nationality, information, education, and social-economic status. Cantril finds no significant difference in the opinions of the well informed and the poorly informed if education and social-economic factors are held constant. The opinions of both informed and uninformed are affected in the same direction by the trend of events.

Education is most influential with respect to opinions involving estimates of quantitative relationships or insight into the possible implications or consequences of an event. The well-educated are also more likely to stand solidly together where opinion involves common values or standards of judgment which are inculcated in the educational system. On issues involving proposals for action or immediate political decisions, Cantril concludes that economic status is decidedly the most important factor in determining opinion. Education and information are effective chiefly in enabling the groups at the higher economic levels to have a clearer conception of where their self-interest lies. Individuals who are able to interpret the trend of events or legislative proposals in terms of self-interest will mainly be found on the side on which they think their interest lies. It seems likely that this tendency extends through all economic levels, limited only by capacity for interpretation. It is also probable, although not pointed out by Cantril, that people who do not trust their own analysis of the issues will usually take their cues from someone who has a definite opinion. This may work in either a positive or a negative direc-

tion. An individual of this type may line up with someone with similar interests whose judgment he trusts or whose favor he is seeking. By contrast, he may line up against someone whose interests are felt to be antagonistic, assuming that he can safely oppose what such an individual favors. Thus it appears consistent with Cantril's general findings to assume that the structure of public opinion is largely determined either by direct awareness of self-interest or indirectly by the reactions of others to those with fixed judgments as to where their interests lie. If these findings can be taken as providing a fundamental picture of the formation and structure of public opinion, they constitute a basic starting point in foreseeing the future development and enhanced social usefulness of the polls. Cantril states this central issue as follows in contrasting the general views of the lower and upper economic groups:

The opinions of the poor group can be generally described as a desire for growth through change in the social structure; the opinions of the upper group as a desire for growth without change in the social structure.

PUBLIC OPINION RESEARCH AND POLITICAL SCIENCE

The National Opinion Research Center, founded in 1941 under a grant from the Field Foundation, is similar to the Office of Public Opinion Research in the general scope of its activities. A close working relationship is maintained in part by the fact that Cantril of the Princeton organization is on the Board of Trustees of the National Opinion Research Center. However, the latter organization apparently represents the viewpoint of political science rather than that of social psychology.

One of the stated purposes of the Opinion Research Center is to maintain a nation-wide staff of interviewers who are available to undertake opinion surveys by various government agencies. Such work has been done for the Office of War Information, the Office of Price Administration, the War Production Board, and state and local government units as well as other federal agencies. Many studies made by the Research Center also reflect a primary interest in the operation of democratic political machinery and what the polls can contribute toward improvement. One study deals with the nonvoter and the special character of his social attitudes and political opinions. Another analyzes the extent and direction of the influence exerted by the Political Action Committee of the C.I.O. A third undertakes to evaluate public attitudes toward politics as a field of activity and toward the quality of individuals operating as professional politicians.

In all these studies the National Opinion Research Center would appear to be more concerned with the harnessing of public opinion for effective demo-

cratic government than with the problems of social psychology underlying the formation of public opinion and its measurement. The study of the nonvoter shows striking and significant differences in relation to economic level. The percentage not taking the trouble to go to the polls increases steadily as income decreases. There is reason to believe that this situation may have been changed considerably by the drive of such groups as the Political Action Committee to get voters to the polls. Calvin Schmid of the University of Washington made a similar check on voting in Seattle, where labor has long been highly organized. He did not find any significant differences in the percentage of eligible persons voting at each economic level.

A general conclusion to be drawn from the study by the Denver organization is that next to income the most important influence is the individual's attitude toward the way in which public affairs are conducted. Those who in general have confidence in politicians and government officials, or who feel that their interests are being properly safeguarded, are the ones who fail to vote. On the other hand, minority groups who are less confident on these points are more inclined to vote when given the opportunity. People of Jewish faith rank high in the percentage going to the polls, with Catholics next. The percentage of Negro voters is low, but this is largely due to disfranchisement in the South.

Harry Field, the director of the National Opinion Research Center, concludes that nonvoters as a group are more opposed to change than voters. When questioned on such points, they react more sharply to any criticism of the American form of government and generally oppose any change in the Constitution. Nonvoters are more isolationist in their views on world problems. More nonvoters than voters are without opinions on the issues of the day. In summary, it appears that nonvoters generally uphold the status quo, appear confident that in general a conservative viewpoint is held by those who are running the government, and see no reason why they should take the trouble to study the issues or go to the polls as long as this is the case.

The effort of the National Opinion Research Center to evaluate the influence of the Political Action Committee of the C.I.O. raises some interesting points, but can hardly be judged to have reached reliable conclusions. In a nationwide poll completed just before the 1944 election, every person who said he was more likely to vote for Roosevelt because of the support of the Political Action Committee was offset by eleven who said they were less likely to vote for him as a result of this support. There can be little question that political activity often consolidates opinion on the opposite side as well as attracting votes to the side that is favored. A campaign speech by a presidential candidate can alienate as well as attract voters if it is excessively partisan or injudicious. Field seems to infer, however, that the Political Action Committee was a

liability rather than an asset to the Roosevelt campaign, a conclusion scarcely justified on the evidence. It is likely that the majority of the people expressing disapproval of C.I.O. support had no intention of voting for Roosevelt anyway. The Political Action Committee was not particularly endeavoring to influence opinion, but devoting itself chiefly to the task of getting to the polls people who shared its opinions. If its activities were not all net gain for Roosevelt, the offset was probably in the form of voters of opposing views who were moved to go to the polls through the publicity given to the C.I.O. activities.

The survey on public attitudes toward politics and politicians indicates a general feeling that politics is not an attractive or even an honorable calling. Only two out of every ten Americans would like to see their sons embark upon a political career, and seven would be definitely opposed. Opposition runs highest among people with college educations, people in the upper income level, and members of professional, business, and white-collar groups. A majority think that it is almost impossible for a man to stay honest if he goes into politics. An even larger proportion are dissatisfied with the way in which most officeholders are doing their jobs. In all these studies of non-voters, of political action groups, and of politics as a profession, there is the common objective of bringing about improvement in political processes. The National Opinion Research Center has established a distinctive position for itself in this direction, along with its activities of taking polls on public issues and having some share in the continuous improvement of techniques for public opinion research.

HOW A PUBLIC OPINION SURVEY IS MADE

The best way of illustrating how all the factors in a public opinion survey are fitted together is to take one poll on a given subject and trace it through from beginning to end. For this purpose a public opinion poll made by the National Opinion Research Center has been selected. This poll attempted to measure on a national scale the opinions of the American public with respect to world organization.

The questionnaire. The National Opinion Research Center selects the subjects for its polls through consultation with its Board of Trustees and other experts in the field. The present subject was chosen for a major project during 1943 because of the great amount of public interest and discussion current at that time. The immediate stimulus seems to have been the Fulbright resolution which had been passed by an overwhelming majority of the House. That broad expression of national policy and intention read as follows:

That the Congress hereby expresses itself as favoring the creation of appropriate international machinery with power adequate to establish and maintain a just and lasting peace among the nations of the world, and as favoring participation by the United States therein through its constitutional processes.

When the subject had been decided upon, members of the staff of the National Opinion Research Center collaborated in drawing up a questionnaire. Each question was carefully pretested with various types of respondents. Since the type of question chosen was of the relatively simple dichotomous type, the main purpose of pretesting was to secure clarity of wording. An effort was made to eliminate all possibility of ambiguous interpretation and arrive at questions that could be understood by respondents everywhere and which were neutral enough to reflect all shades of opinion accurately.

The field work. Detailed instructions for interviewing were prepared with the aim of ensuring uniform interviewing procedure. The Center has a staff of two hundred trained interviewers throughout the country and generally sends a quota of interviews to every member of this field organization. Along with the instructions and the interview forms, each interviewer receives a quota sheet specifying the controls he is to use in selecting his respondents.

In the present instance interviewers were instructed to cover people twenty-one years of age and over and to distribute them according to a certain pattern by sex, age, economic level, and geographical location. The geographical distribution started with a proportionate allocation to each of the nine census divisions of the country, with further controls by size of city, and by farmers and nonfarmers in rural areas. The distribution by age and sex took account of the changes brought about by young men being pulled out of their home communities for military service. Thus, instead of the sex ratio of approximately fifty-fifty as shown for persons twenty-one years and over in the 1940 Census, estimates were made concerning the present population, which resulted in 52.9 per cent of the interviews being made with women. Similarly the age distribution for 1943 was estimated so that 53.4 per cent of the interviews were made with persons forty and over, a higher proportion than in the normal civilian population. Negro respondents were included according to their relative positions in the population in each census division and according to the urban, rural, nonfarm, and farm classifications in each level.

Economic status was designated in terms of four standards-of-living categories. The two upper groups are thrown together for purposes of tabulation so that three separate figures are reported in the findings. This division by economic status amounts roughly to classifying the total population into an upper quarter, a middle half, and a lower quarter. These economic status

classifications are identified in the field largely by observation of comforts or luxuries present in the home.

In this survey the National Opinion Research Center used 2,448 interviews. This number would be sufficient 997 times in 1,000 to keep the error in the results within limits of 3 per cent on questions that divide evenly. The number is high enough to provide a somewhat lower, but still satisfactory, degree of accuracy on questions showing a preponderant majority in one direction.

Analysis of results. The material received was subject to machine tabulation. Each item of information on a questionnaire was translated into a code number and all the information then transferred to punched cards, one card for each questionnaire. The cards were then put through the mechanical sorter a number of times. Counts were secured showing the division of opinion, not only for the sample as a whole but for each of the population subgroups, such as male and female, youth and age, poor and prosperous, farmers and city people, and variations in educational attainment from grade school only to college graduates. The division of opinion along each of these lines of variation in the population were then expressed as percentages as the easiest form for reporting the results.

Results. The first question asked was "If a union of nations is formed after the war, do you think it would be a good idea or a bad idea for the United States to join it?" This question had been asked on three previous occasions reaching back over a period of years, the first poll having been made in September, 1942. The percentage liking the idea of a union of nations steadily increased over the period and the final figure was 81 per cent.

Other questions were asked in the poll of September, 1943, to measure attitudes on somewhat more specific proposals related to the general proposition. First was the question "Do you think a union of nations would have the best chance of working if all countries were members, or if just some of the countries were members?" Here the response showed 83 per cent feeling that all countries should be members, or approximately the same result as on the main question. However, the next point shows a sharp change. Respondents were asked "After the war do you think Germany, Italy, and Japan should be members of the union?" Only 54 per cent said "yes." People often back down when confronted with the specific implications of their general opinions.

Another section of the questionnaire dealt with the powers that might be given to a world union. Replying to a series of questions,

69 per cent said that international disputes should be settled by the organization of nations rather than solely by the parties to the dispute;

76 per cent favored giving the union of nations authority to make laws concerning international problems;

and of those who favored this legislative authority, 68 per cent favored an international police force to make all nations obey the laws.

Questions on the control of armaments again reveal inconsistencies in public attitudes. Respondents were asked whether the union of nations should decide how big the armed forces of each country should be. Replies in the affirmative constituted 77 per cent. A somewhat smaller group, or 70 per cent, were willing to have this control applied to the United States. Almost completely counter to these expressions was the fact that 89 per cent thought it would be necessary for the United States to maintain a larger army, navy, and air force than before the war. By contrast, only 22 per cent felt that every nation should be permitted to decide for itself how large its armaments should be in the postwar period.

This poll on world organization asked many other questions on such subjects as regulation of trade, prevention of wars, other functions of the world union, and the willingness of Americans to make specific sacrifices for the sake of a world union. Enough of the results have been summarized, however, to illustrate a fundamental principle: logical inconsistency is an inherent characteristic of public opinion, or of nearly any individual's opinions, for that matter. Socrates demonstrated long ago that the average human being seldom examines the relationships of beliefs he might hold in separate fields or to face the consequences of a particular belief with respect to proposed action.

THE OUTLOOK FOR THE PUBLIC OPINION POLLS

In this final section an attempt will be made to point out some of the hazards affecting the future of the public opinion polls and to evaluate the opportunities for further constructive development. The discussion of the hazards can be crystallized by reference to the Congressional hearings on the American Institute of Public Opinion. The very fact that the hearing should have been conducted dramatizes the ultimate threat of public regulation. It is the subject matter of the investigation rather than its implications that will be examined here.

The committee was principally concerned with any evidence of intention to influence the election results by distorting the poll results. The committee of technical experts which reviewed the evidence commented primarily on three aspects of Gallup's operation: sampling methods, adjustments made in field returns, and the publication policy of the Institute.

The question of whether the manipulation of poll figures can influence the results of the election is tied up with the theoretical question of whether there is a bandwagon vote. In other words, is there an important segment of the electorate which is influenced to vote by the apparent trend of public

preference? Gallup contended that there is nothing to the theory. By way of proof he said that he had never found any evidence of it in his poll results. He further stated that if that were the case the candidate who received the majority preference in the first poll would improve his position throughout the campaign. Actually the poll results show many exceptions to such a tendency.

Certainly the bandwagon factor could never cause any president to be elected by a unanimous vote. However, it would need to influence only 1 or 2 per cent of the electorate to swing the election if the candidates were only three or four percentage points apart. Also Gallup's theory of how the bandwagon vote would operate if there were such a vote is subject to some refinement. He states that it could only help the candidate in the lead. Might it not work the other way if the first ballot showed that a candidate had a chance when his supporters had been inclined to believe that his candidacy was hopeless?

That is apparently the type of situation which the committee had in mind in calling Gallup to testify. The two-point reduction applied to a succession of the poll reports in the fall of 1944 made it appear that the outcome was a tossup between Dewey and Roosevelt. That might have offered sufficient encouragement to Dewey supporters to account for the Dewey trend which was reported. The trend amounted to only two points, and half of that, according to the technical committee, was accounted for by changes in method of tabulation during the course of the campaign.

A main point of criticism by the technical committee concerned sampling methods. This must have seemed like *lèse majesté* to the Gallup organization, since they had so successfully identified themselves in the popular mind with the whole idea of scientific sampling. The committee pointed out the weaknesses of the system of quota control used by Gallup and the other poll takers and suggested by implication that the system of area control recently developed by the Census Bureau should be adopted.

The last word has not yet been said on the relative merits of these two systems of sampling control.

Accuracy of the poll results is already high enough that the greatly increased cost of the larger sample would scarcely be justified. There are many other factors in the possible improvement of the polls both as measurement and as prediction. Progress in those lines should be sought concurrently with such improvement in sampling as is consistent with the budgets normally available for marketing and public opinion research.

The criticism of the adjustments made in the actual polling figures seems to have greater validity. The technical experts point out that prediction of turnout is as difficult as the measurement of political attitudes itself. Hence, it is hardly to be dismissed by such an arbitrary procedure as the two-point deduc-

tion from the Roosevelt returns carried through the campaign up until the last ballot. The use of this figure as a flat correction state by state also is criticized. One possibility which might at least have been considered was that of applying a smaller correction in the states where the margin was close and which might have a better than average turnout for that very reason. The apparent showing in these sharply divided states would be most affected by the arbitrary adjustment and a number of them were shifted thereby from the Roosevelt to the Dewey column. Returning to the discussion of the bandwagon theory, it might be possible to influence the election as a whole by giving Dewey a specious lead in doubtful states and thus stirring up greater activity for him at these crucial points.

Finally the technical experts criticized the publication policy, which resulted in presenting the adjusted figures without adequate safeguard against creating the impression that these were the actual polling results. This point has been covered in the earlier discussion of Gallup's reluctance to make a clear choice between the viewpoints of measurement and of prediction. As to an improvement over present practice the technical committee suggests that each release might contain a statement that various adjustments had been made in the figures and that a technical bulletin could be obtained on request explaining all the adjustments.

Speaking from the viewpoint of market research and the other aspects of social science which have a stake in the integrity of the polls, what is believed to be a much better suggestion is offered here. That is that the actual polling results be presented without adjustment in each release and that they be clearly labeled as representing only an attempt to measure the attitudes of a representative cross section. Then when Gallup chooses to make a personal prediction, whether it is done only near the end of the campaign or several times during its course, let him explain that he has introduced other weights to allow as best he may for factors other than expressed attitudes. This would appear to be the best way of reconciling the conflicting requirements of measurement and prediction, recognizing the difficulty of getting the average reader to distinguish clearly between the two. This procedure might greatly increase the journalistic interest in the polls, particularly if Gallup would issue a good-natured challenge to his readers to outguess him, and offered to send them the additional data he used in his own forecast if they wished to make a serious attempt.

As to the constructive future of the polls, the possibility that a bandwagon factor may exist should not be considered as a fatal blow. It should be possible to defend that feature frankly if the material is collected as scientifically as possible and presented fairly. The individual is entitled to whatever information is available in making up his mind. If he is one of those who believe that

the majority is likely to be right, then he is entitled to know how most of his fellows are going to vote if he has any way of finding out. A worker still has a right to ask his boss how he is going to vote even though the right of the boss to tell him without solicitation has now been limited by law. A wife can and typically does ask her husband how he is going to vote. The bandwagon factor apparently works quite effectively in the small universe to judge by the small number of cases in which husband and wife vote differently.

There is no grave danger in the small weight that may be carried nationally by the bandwagon vote, remembering that under our system electoral votes are determined state by state. There are plenty of people with sufficient stake in the local outcome to put up a good fight regardless of the national forecasts. Certainly good practice by the poll takers should avoid every suspicion that they are trying to create a spurious bandwagon influence in doubtful states by manipulating poll results. Otherwise there is nothing particularly objectionable to having the majority of people still expressing the same opinion at the end of the campaign as at the beginning, provided the indicated choice at the beginning was as scientifically accurate as possible. Nothing objectionable except making it painfully clear, as many already suspected anyway, that the influence of the campaign upon the result is usually negligible.

One man who should be very much interested in a correct estimate of majority opinion is the individual legislator or public official. Under our democratic system it is presumed that he represents the will of the people. He may sometimes feel that in his superior knowledge of the issue he is justified in going against public opinion. Nevertheless, it is well for him to know what it is. With an authentic public opinion poll before him he need not be so jumpy every time some minority pressure group threatens to throw him out of office if he refuses to do its will.

To accomplish this type of social benefit the local activities in public opinion polls need to be greatly increased. Thus the polls may come back to the local newspapers who started them years ago. The Philadelphia *Evening Bulletin* has recently established such a local polling operation and has already demonstrated that constructive results can be obtained. During the summer of 1944 the city of Philadelphia suffered a streetcar strike because of the upgrading of Negro employees. A poll conducted by the *Bulletin* showed that a majority of the citizens favored the cause of the Negro worker. There is good reason to believe that the publication of this result helped to relieve racial tension by moderating the Negroes' feeling of constituting a persecuted minority.

Many civic bodies in Philadelphia have also found the polls helpful. Some have been agreeably surprised at the strength of the support for a worthy project. Others, who may be equally sincere in advocating less popular measures, are helped by getting a more realistic appraisal of the size of the

job ahead of them. Retailers have profited from some of the polls taken on the degree of interest in certain holidays and on other matters.

At its best the public opinion poll completes the flow of two-way communication between the people and their leaders. The public official can speak to the people through the press or over the radio. The polls give the people the opportunity to talk back in clear and unmistakable terms. The ultimate ideal is to have them so well organized locally as to form a channel of discussion which can substitute effectively for the old-fashioned town meeting.

As a final speculation one might consider the possibility of having a newspaper's editorial page reflect only the majority opinion of its readers. The foundation of this approach would be a continuous polling operation. Leading citizens or national authorities might then be asked to write editorials presenting the arguments for the majority opinion. Key facts or statistics useful to people in considering the issue might also be presented from time to time. Whatever the course the local development takes, it is certain that it will expand and will greatly strengthen a contribution to social research that is already so well established on the national scale.

Index

- Advertising analysis
 see Advertising Research
- Advertising
 analysis, 66-70
 copy development, 96-117
 copy testing, 108, 118-136
 controlled opinion tests, 119, 128-129
 inquiry and coupon tests, 119, 130-133
 recall tests, 119, 123-128
 recognition and identification tests, 119, 120-123
 sales tests, 119, 133-135
 research, 3
- Advertising Research Foundation, 117, 136
- Alderson, Wroe, 4, 5, 289-309
- Alderson, Wroe, Simon and Sessions, 289-309
- Allen, G. H., 151-155
- American Institute of Public Opinion
 see Gallup, G. H.
- Analysis of Results
 see Results, Analysis of
- Arnheim, Rudolph, 145, 150
- Area sampling
 see Sampling, area
- Atlantic Monthly*, 211
- Attitude
 consumer, 12
 employee
 see Employee's morale
 farmers
 see Farmers' Attitudes
- Audit Bureau of Circulation, 209
- Avery, S. W., 88
- Barton, S. G., 3, 58-70, 181-187
- Beville, H. M., 139, 150
- Blankenship, A. B., 1-7
- Bordon, N. H., 128, 136
- Brand Barometer
 see Psychological Corporation
- Branded goods, 3, 5
 movement of, 58-70, 71-85
- Broadcasting Magazine*, 158
- Brown, T. H., 215, 261
- Brunsmann, H. G., 255, 258
- Buying habits, 44
- Campbell, A. A., 4, 274-285
- Cantril, Hadley, 107, 117, 150, 296-300
- Caples, John, 130, 136
- Carvl, F. R., 18
- Census Surveys
 see U. S. Government
- Channels of marketing
 see Marketing, channels of
- Coincidental method, 151-155, 156-171, 184
- Collier's*, 122, 216
- Columbia Broadcasting System, 183, 185, 186
- Commercial Surveys
 see Surveys, commercial
- Committee for Economic Development, 43-57, 232-239
- Conclusions
 see Results, conclusions from
- Confusion control, 215-219
- Consumer attitude
 see Attitude, consumer
- Consumer Jury Test
 see Advertising Research—Controlled opinion tests
- Consumer panel
 see Panel
- Consumer purchases
 see Branded goods, movement of
- Controlled opinion tests
 see Advertising Research
- Controlled sampling
 see Sampling, controlled
- Cooperative Analysis of Broadcasting, 3, 4, 151-155, 156
- Copy development
 see Advertising research
- Corbin, H. H., 85
- Coronet*, 5
- Country Gentleman*, 4, 189, 190, 192, 195, 198, 202
- Coutant, Frank, 133, 136

- Crossley, A. M., 4, 215, 261, 289, 290, 292-294
 Curtis, Alberta, 140, 150
 Curtis Publishing Co., 21, 188-208
- Dealer surveys, 232-239
 Dedrick, C. L., 251, 258
 Demand, analysis of, 33
 Deming, W. E., 255, 258
 Department of Agriculture
 see U. S. Government
 Depth interviews
 see Interviews, open-ended
 Dewey, Thomas, 293-295
 Dichter, Ernest, 103, 106, 117
 DuBois, Cornelius, 3, 209-221, 261
 DuPont, E. L., de Nemours & Co. Inc., 11-28
- Editing of questionnaires
 see Questionnaire, editing
 Editorial research, 188-208
 Efficiency of marketing
 see Marketing, Efficiency of
 Employees' morale, 4, 222-231
Evening Bulletin, 308
 Experimental studies, 106-108
- Factory management and maintenance, 93
 Falk, A. T., 136
Farm Journal, 289
 Farmers
 attitudes, 275-276
 behavior, 277-278
 information, 277
 Federal Advertising Agency, 86-95
 Field, Harry, 301
 Field work
 respondent co-operation, 297-298
 supervision, 14, 19-20
 Firth, L. E., 136
 Fiske, Marjorie, 3, 137-150
 Fleiss, Marjorie, 146, 148, 150
Fortune Magazine, 4, 5, 290
 Frankel, L. R., 252, 258
 Franzen, Raymond, 215, 261
 Freiberg, A. D., 3, 44, 71-85, 118-136
- Gallup, G. H., 4, 5, 215, 261, 289, 290, 293-296, 298, 305-307
 Garrett, P. W., 87, 95
 Gaudet, Hazel, 150
 Giellerup, S. H., 133-134, 136
Good Housekeeping, 211
 Government research
 see U. S. Government Surveys
- Grant, H. B., 4, 120
 Grimshaw, Austin, 261
- Hallonquist, T., 150
 Hansen, M. H., 251-259
 Haring, C. E., 3, 43-57, 232-239
Harpers' 211
 Hauser, P. M., 4, 251-259
 Herzog, Herta, 144, 145, 150
 Hilgard, E. R., 4, 253, 258, 259-273
 Hooper, C. E., 3, 4, 156-171
 Hooperatings, 168-171
 Hooper-Holmes Bureau, 46, 234, 235, 238
 Houser Associates, 222-231
 Houser, J. D., 105, 117
 Hurwitz, W. N., 252, 258
- Industrial relations
 see Public relations
 Industrial research and survey, 43-57
 Comparison with consumer research, 45-46
 Industrial Surveys Co., 3, 4-5, 58-70, 135, 181-187
 Inquiry and coupon tests
 see Advertising Research
 Instructions
 see Interviewers, training of
 Interpretation
 see Results, interpretation of
 Interviewers
 bias, 298-299
 rapport, 298-299
 selection of, 19
 training of, 19, 47-48, 78, 195-198, 303
 see also Field Work
 Interviews
 follow-up, 78
 open-ended, 103-104, 278-279
 personal, 18-20
- Jenkins, J. G., 83-85, 103, 117
 Jessen, R. J., 279
- Karslake, J. S., 115, 117
 Katz, Daniel, 4, 241-250
 Kelly, Nelle, 172-178
 Kettering, C. F., 46, 47
 King, A. J., 279
 Knox Reeves Advertising, Inc., 96-117
 Kolstad, Arthur, 4, 222-231
- Lazarsfeld, P. F., 3, 103, 117, 137-150, 215, 261
 Lehman, A. W., 151-155
 Lewis, J. L., 88

- Liberty*, 122, 216
Life Magazine, 3, 212-222
 Likert, Rensis, 21, 275
 Link, H. C., 76, 85, 89, 106, 117, 125, 136
 Listener Diary Surveys, 185-186
Literary Digest, 60, 289, 291
 Lockley, L. C., 2, 6, 11-28
 Lorge, Irving, 85
 Lovekin, O. S., 128, 136
 Lucas, D. B., 21, 121-123, 136
 Ludeke, H. C., 3, 188-208
- Macfadden Publications, 92
 Magazine Audience Group, 215-221
 Mail questionnaire
 accuracy of, 17
 advantages of, 17
 cost of, 17
 disadvantages of, 16
 Mangus, A. R., 195
 Market
 description, 2, 11-28
 research, 43-44
 study, 33-34
 Marketing
 channels, 13, 32, 232
 efficiency, 13
 Market Research Co., 29-42
 McCandless, Boyd, 141, 150
 Media measurement, 3, 66
 Meyers, Bernard, 137
 Meyrowitz, Alvin, 140, 150
 Morale, employees'
 see Employees' morale
 Murphy, M. J., 121-122, 136
 Myers, H. B., 252, 258
- National Analysts, Inc., 1-7
 National Consumer Panel, 60-70, 135
 National Opinion Research Center, 4, 245,
 290, 295, 300-304
 New uses for products
 see Products, new uses for
 Newhall, S. M., 128, 136
News Week, 293
 Nielsen, A. C. Co., 135, 186
 Noyes, C. E., 4, 259-273
- Observation method, 21-22, 106
 Office of Public Opinion Research, 290, 295,
 300
 Office of Radio Research, 137-150
 Office of War Information
 see U. S. Government Surveys
- "Open Question"
 see Interviews, open-ended
 Opinion Research Corporation, 93
- Panel, 17-18, 58-70, 90, 148
 accuracy of, 65
 application of, 65-70
 limitations of, 187
 operation of, 60-64
 radio, 181-187
 values of, 18, 60-64, 184
 Payne, S. L., 253, 258, 273
 Personal Interview Roster Method, 172-178
 Personal interviews
 see Interviews, personal
 Philadelphia *Evening Bulletin*
 see *Evening Bulletin*
 Phrasing of questions
 see Questions, phrasing of
 Polls
 see Public opinion polls
 Pre-test, 23
 Problem
 coverage of, 100, 101
 statement of 14-15
 Product
 acceptance, 37
 design, 36
 research, 3, 12, 29-42, 44, 46-57, 188
 Products, new uses for, 13
 Program analyzer, 143-144
 Psychological Corporation 5, 78, 119, 12,
 124, 126, 135, 136
 Brand Barometer, 5, 71-85, 89-90, 135
 Public opinion polls, 4, 289-309
 Public relations, 44, 86-95
 Pulse, The, 3, 4, 5, 172-178
 Purdue Eye Camera, 115-116
 Purposive sampling
 see Sampling, purposive
- Questionnaire
 coding, 80
 editing of, 14, 80
 formulation of, 14, 22-23, 48-55, 102-10
 mail, *see* Mail questionnaire
 pre-coding, 23-24
 reproduction, 78
 tabulations of, 23
 Questions
 objective, 104-105
 phrasing of, 22, 55, 102, 107-108, 22;
 296-297

- Radio research, 137-150, 151-155, 156-171, 172-178, 181-187
 audience measurement, 151-155, 156-171, 172-178, 181-187
 commercials, 108, 110, 118, 146-147
 content analysis, 142
 monitoring, 141
 objectives, 137-39
 techniques, 139-150
- Ratio sampling
see Sampling ratio
- Reader's Digest*, 5
- Readership surveys, 3, 44, 188-208, 209-221
 methods of, 211-222
 problems of, 210-211
- Recall method, 184
- Recall tests
see Advertising research
- Recognition tests
see Advertising research
- Report, 27-28
- Respondent cooperation
see Field Work
- Results
 analysis of, 14, 27-28, 91, 99-100, 104-105
 conclusions from, 100
 interpretation of, 80
 sponsor, 57
- Retail surveys
see Dealer surveys
- Robinson, W. S., 140, 150
- Roosevelt, F. D., 293-295, 301-302, 307
- Root, A. R., 110
- Roper, Elmo, 215, 261, 289, 291, 294
- Roslow, Sydney, 172-178
- Rudolph, H. J., 136
- Sales tests
see Advertising research
- Sample
 selection of, 14
 size, 24-25, 93, 280
- Sampling
 area, 25-26, 252-254, 279-280, 295
 controlled, 25-26, 77, 89-90, 93, 291, 303-304
 purposive—*see* Controlled
 quota—*see* Controlled
 random, 291
 ratio—*see* Controlled
 selection—*see* Controlled
 stratified—*see* Controlled
- Saturday Evening Post*, 5, 121-123, 124, 211, 216
- Schmid, Calvin, 300
- Selective sampling
see Sampling, selective
- Size of sample
see Sample, size
- Smith, C. H., 150
- Smith, E., 106, 117
- Smith, E. R., 92, 95
- Smith, Kate, 138
- Split-run technique
see Advertising research—inquiry tests
- Stanton, Frank, 143
- Starch, 4, 120, 128, 136
- Stephan, F. F., 255, 258
- Stock, J. S., 252, 258
- Stratified sampling
see Sampling, stratified
- Strong, E. K., 105, 117, 128, 136
- Suchman, E. A., 106, 117, 150
- Surveys
 commercial, 2
- Tabulations
see Results, analysis of
- Telephone interviews
see Interviews, telephone
- Test
see Pre-test
- Test cities, 133
- Turnbull, 107
- U. S. Government Surveys, 44-45
 Army, 4
 Bureau of Agricultural Economics, 274-285
 Bureau of the Census, 5, 251-259
 Office of War Information, 4, 241-250
 War Production Board, 259-273
- War Production Board
see U. S. Government
- Webb, John, 252, 258
- Welch, A. C., 3, 96-117
- Welles, Orson, 138
- White, Matilda, 3, 29-42
- White, Percival, 3, 29-42
- Wilks, S. S., 215, 261
- Willmark Research Corporation, 5
- Wulfbeck, W. H., 3, 86-95