Proceedings of the Sawtooth Software Conference on Perceptual Mapping, Conjoint Analysis, and Computer Interviewing

1987

Foreword

These are proceedings of the Sawtooth Software Conference at Sun Valley, Idaho in March, 1987. The conference topics were Computer Interviewing, Perceptual Mapping, and Conjoint Analysis. There were approximately 250 participants.

The speakers were not asked to provide written versions of their talks; rather, their remarks were recorded and transcribed. The informal nature of some of the papers adds to their freshness.

Anyone who has seen an unedited transcription of his spoken words is likely to be distressed. Even the most fluent of us seem to speak in incomplete sentences, dart off on tangents, and mix up our cases and tenses. Most of the editing of the transcriptions was done by Carol Potera, with the assistance of the authors. The editing and production process required about four months. Next year we will try to abbreviate this by asking our speakers to provide written versions at the time of the conference.

We think the information in this volume will be of use to a wide range of readers. We thank our speakers for their thoughtful presentations, and we invite our readers to comment. In particular, we will be grateful for any suggestions about how to make future conferences even more successful.

Richard M. Johnson
July 20, 1987

Contents

COMPUTER INTERVIEWING

Historical Perspectives and the Future of Computer Interviewing
Lawrence Dandurand, University of Nevada.............. 1

Doing Traditional Research By Computer: What Has Been Learned So Far

Long Self-Administered Questionnaires
Lesley Bahner, POPULUS, Inc............................ 11

Political Polling
Brent Stahl, MORI Research 23

A Client's Experience With Computer Interviewing
David Griscavage, Pepsico, Inc......................... 33

Telephone Interviewing
Richard Miller, Consumer Pulse, Inc.................... 39

Children's Research and Other Issues Related to the Use of Computerized Interviewing
Ira W. Goodman, J.M.R. Marketing Services, Inc......... 47

New Horizons - Taking Computers Where They Haven't Been Before

Computer Surveys By Mail
Harris Goldstein, Trade-Off Marketing Services......... 55

Complex Interviews With Laptop Computers
Joel Gottfried, National Analysts...................... 61

Computers and Hard-To-Interview Respondents
William Tooley, IMR Systems, Inc....................... 69

Who Should Do What? - Field vs. Supplier vs. Client

Audrey Bowen, General Foods............................ 73
David Griscavage, Pepsico, Inc......................... 77
David Santee, Hallmark Cards, Inc...................... 79
Peter Honig, Peter Honig Associates.................... 83
Elizabeth Bradley, National Analysts................... 87
Richard Miller, Consumer Pulse, Inc.................... 91
Bernadette Schleis, Interactive Network................ 97

Ci2 in the University

Implementing a Ci2 System Multiuser License in a State University
Arthur Saltzman, California State University..........101

Ci2 in the University and Teaching Future Researchers
Lawrence Dandurand, University of Nevada...............117

PERCEPTUAL MAPPING

Perceptual Mapping: Its Origins, Methods, and Prospects
Allan D. Shocker, University of Washington.............121

Adaptive Perceptual Mapping
Richard M. Johnson, Sawtooth Software, Inc.............143

How to Sell Perceptual Mapping

How to Sell Perceptual Mapping
Michael H. Baumgardner, Burke Marketing Research.......159

How to Sell Large Market Research Projects Internally Within Client Companies
Edward (Ted) Evans, Chevron Chemical Company...........169

How to Design A Study

How to Design A Study
Betty A. Sproule, Hewlett Packard......................173

How to Design A Perceptual Mapping Study
Bill McLauchlan, Burke Marketing Research..............179

Analysis and Interpretation of Results

Analyzing and Interpreting Perceptual Mapping Data
Paul N. Ries, Procter & Gamble.........................189

Analysis and Interpretation of Results
Paul Hase, Hase/Schannen Research......................199

Presenting Results

Presenting Results
Bruce Morrison, General Electric.......................219

Presenting the Results of Perceptual Mapping Studies
John A. Fiedler, POPULUS, Inc..........................225

CONJOINT ANALYSIS

Conjoint Analysis: How We Got Here and Where We Are
Joel Huber, Duke University....237

Adaptive Conjoint Analysis
Richard M. Johnson, Sawtooth Software, Inc....253

How to Sell Conjoint Analysis

How to Sell Conjoint Analysis
Verne B. Churchill, Market Facts, Inc....267

How to Sell Conjoint Analysis
Vincent P. Vaccarelli, Xerox Corporation....275

How to Design A Study

How to Design A Conjoint Study
James J. Tumbusch, Procter & Gamble....283

Designing a Conjoint Analysis Study
Richard D. Smallwood, Applied Decision Analysis, Inc...289

Analysis and Interpretation of Results

Conjoint Experiments : Analysis and Interpretation of Results
Marshall G. Greenberg, National Analysts....297

Analysis and Interpretation of Conjoint Results
Donald Marshall, Smith Kline & French Laboratories....307

Presenting Results

Presenting Results
Peter B. Bogda, M/A/R/C....317

Presenting Results
Marc Prensky, The Boston Consulting Group....327

HISTORICAL PERSPECTIVES AND THE FUTURE OF COMPUTER INTERVIEWING

by
Lawrence Dandurand
Department of Marketing
University of Nevada
Las Vegas, NV

Tull and Hawkins in Marketing Research (1987, p. 104) state that "computer interviewing is beginning to gain acceptance." Parasuraman (1986, p. 374) indicates that "recent technological advances are leading to increased use of computerized interviewing, such as centralized telephone interviewing with the aid of CRT terminals and questionnaires filled out by respondents sitting at computer keyboards in shopping malls." Gorves and Mathiorvets (1984, p. 386) believe that computer-assisted telephone interviewing (CATI) will eventually dominate telephone interviewing.

What is computer-assisted interviewing? How does it relate to the marketing research process? What has been the history of computer-assisted interviewing? Which technological advances are relevant to the development of computer-assisted interviewing? What social trends are affecting the development process? What is the future of computer-assisted interviewing? The purpose of this paper is to answer these questions and to suggest directions for future study.

Computer-Assisted Interviewing and the Research Process

Computer-assisted interviewing is asking and answering questions in a computer-based system. The computer presents the questions to interviewers or respondents via CRT's (Cathode Ray Tubes), TV screens, print-data terminals, or PC monitors. Interviewers or respondents enter answers using a keyboard, keypad, a light pen, or by finger-touch. Computer interviewing can be employed in personal, telephone, mail, and combination research designs.

Computer-assisted interviewing is a high-tech research tool that can be integrated into the research process. It is designed to facilitate the research process. It is not designed to replace phases of the research process, such as questionnaire design, field methodology, or data analysis.

Historical Perspectives

Computer-assisted interviewing has been possible for 25 years. It was made possible by the introduction of interactive remote terminals and interactive programming. IBM introduced its interactive remote terminal, the SABRE System, in 1962. It introduced the interactive system just two years after it had introduced its popular 1400 Series of mainframe computers. Table 1 shows a history of technological developments affecting the development of computer-assisted interviewing.

TABLE 1

Technological Developments Affecting Computer-Assisted Interviewing

Technological Development	Company	Year
1. Mainframe Computers	IBM	1960
2. Interactive Remote Terminals	IBM	1962
3. WATS	AT&T	1964
4. Interactive CRT Systems	IBM	1964
5. Mini-computers	DEC	1965
6. Telecommunications Networks	Telenet	1975
7. Personal Computers	IBM	1981
8. PC Ci2 Software	Sawtooth	1985
9. PC CATI Software	Sawtooth	1986

Current Role of Computer-Assisted Interviewing

The popularity of computer-assisted interviewing is increasing. It is currently most popular with WATS (Wide Area Telephone Service) and central location-type research (e.g., shopping malls, trade shows, and product clinics). Churchill (1987, p. 254) writes that "partly because of the advantages that accrue with CRT administration of questionnaires, telephone interviews are currently the most popular [of the research techniques] and have experienced the greatest increase in popularity over the last few years among members of the Council of American Research Organizations (CASRO)." Table 2 (Honomichl, 1984) depicts the relative popularity of alternative research techniques from 1981 through 1983.

TABLE 2

Relative Use of Research Techniques

	1983	1982	1981
WATS/CENTRAL TEL. FACILITY	41%	38%	38%
CENTRAL LOCATION/MALL	19%	17%	16%
	---	---	---
SUB-TOTAL	60%	55%	54%
OTHER (PERSONAL, NON-WATS PHONE, MAIL, FOCUS GROUP)	40%	45%	46%
	----	----	----
TOTAL	100%	100%	100%

Table 2 indicates that WATS and central location research constituted 60 percent of the research conducted in 1983. These two techniques were the most likely to incorporate computer-assisted interviewing and they experienced a six percent increase in the two year period between 1981 and 1983.

Advantages of Computer-Assisted Interviewing

The basic advantages of computer-assisted interviewing are bias reduction, error control, time saving, response improvement, and data integration. Bias reduction is achieved, for example, with the ability to randomize section, question, and answer sequence. Error control derives from the ability to build logic into the system; e.g., if a given scale has a range of 1 to 5, any answer outside of this range would automatically be rejected. The time saving arises, for instance, from not having to keypunch or "clean" the data or being able to provide real-time top-of-the-line research results. Response improvement comes from customizing, personalizing, automating, and "autonomizing" the questionnaire electronically. Data integration refers to the ability to link with statistical programs and computer-based marketing models and merge with other data sets. Table 3 lists specific advantages associated with computer interviewing.

TABLE 3

Advantages of Computer-Assisted Interviewing

1. Randomizing Response Choices
2. Checking for Response Consistency
3. Incorporating Complex Skip Patterns
4. Personalizing (Forward Control)
5. Adding New Response Categories
6. Customizing Questionnaires
7. Producing Instantaneous Data Analysis
8. Managing Interviewer Variability
9. Controlling Sample Selection
10. Making Instantaneous Revisions
11. Automating Callback Schedules
12. Accessing Related Information
13. Providing Research Management Reports
14. Conducting Multi-Lingual Interviews
15. Increasing Flexibility (Ease of Revision, Etc.)
16. Reducing Costs (Different Versions, Changes, etc.)
17. Improving Research "Turn-around" Time

Problems Associated With Computer-Assisted Interviewing

Billing (1982, p. 2) says to "go slow, be wary when considering a switch to a computer-assisted interviewing system." He believes that functions such as editing and keypunching will be largely eliminated. However, he states that new functions - such as debugging, programming logic, and computer maintenance - will be created. In short, a high-tech decision will affect the fundamental structure of the research organization.

Of course, there are other problems to consider. For example, executives and research staff members might place too much faith in high-tech computer interviewing to the detriment of good questionnaire design. If the design is bad, the computer will not automatically correct it. Also, self-administered computer interviews usually require typing skills. This is especially true for open-ended questions. A respondent who cannot type has a difficult and frustrating time completing the interview. In addition, there are problems associated with hardware and software costs and capabilities, hardware and software support in a dynamic and turbulent industry, respondents who are unwilling or unable to use computers, and inexperienced field organizations and field interviewers. Table 4 summarizes the problems associated with computer-assisted interviewing.

TABLE 4

Problems Associated with Computer-Assisted Interviewing

1. Debugging Time
2. Programming Logic
3. Computer Maintenance
4. Misplaced Faith
5. Typing Skills
6. Hardware Costs
7. Software Costs
8. Hardware Support
9. Software Support
10. Unwilling/Incapable Respondents
11. Inexperienced Field Organizations
12. Inexperienced Field Interviewers

The Future Role of Computer-Assisted Interviewing

The future role of computer-assisted interviewing will be a function of technological developments, social trends, computer interviewing linkages, and a change in marketing research orientation. All four of these factors indicate continued growth of computer interviewing. Moreover, the nature and practice of computer interviewing will change dramatically over the next ten years.

Technological Developments

Table 5 lists technological developments that will affect the use of computer-assisted interviewing in the research process. For example, a voice response system will make the computer seem more like a telephone. It will help to overcome computer shyness and lack of typing skills. Furthermore, these technological innovations are in the initial stages of product development. Zikmund (1986, p. 221) suggests that this type of high-tech is "allowing for creativity in the modes of data collection."

TABLE 5

Technological Developments Affecting Computer-Assisted Interviewing

1. Personal Computers
2. Lap-top Computers
3. Microprocessors
4. Computer Networks
5. Light Pens
6. Finger-touch Monitors
7. Entry Pads
8. Voice Response Systems
9. Software
10. Cable Television
11. Two-way Cable Television
12. Antenna Dishes
13. Scanners
14. Sensors
15. VCRs
16. Robotics
17. Video Phones
18. Laser Printers
19. Video Disks
20. Telecommunications
21. Videotex Terminals

Social Trends

There are social trends that are favorable to the increased use of computer-assisted interviewing. For example, the spread of personal computers in households and businesses will permit the mailing of interviewer-administered computer assisted interviews to respondents. Table 6 provides a list of these trends.

TABLE 6

Social Trends Affecting Computer-Assisted Interviewing

1. Spread of Personal Computers
2. Widespread Ownership of Telephones
3. Growth of Cable and Satellite Television
4. Acceptance of High-tech Products
5. Difficulty of Door-to-Door Research
6. Shortage of Experienced Interviewers

Computer-Assisted Interviewing Linkages

Environmental factors are evolving that are linked to computer-assisted interviewing and will make computer-assisted interviewing more attractive. In other words, the existence of these factors will pose benefits not only in themselves, but benefits to be achieved by linking with computer-assisted interviewing. In addition, it is expected that these factors or forces will be combined to create additive and synergistic possibilities. For example, as electronic data search becomes less expensive and more readily available, and personal computers diffuse into households and businesses, it will become easier to conduct secondary research and computer-assisted primary research and to merge the results of both.

TABLE 7

Computer-Assisted Interviewing Linkages
Affecting Computer-Assisted Interviewing

1. Data Processing Software
2. Interlocking Databases
3. Electronic Data Search
4. Computer-based Models
5. Expert Systems
6. Artificial Intelligence
7. Automated Offices
8. Management Decision Support Systems
9. Electronic Publishing
10. Information Brokers

Change In Marketing Research Orientation

The orientation of marketing research professionals, with respect to how they perceive the performance of the research function, will change during the next ten years. This change will be the result of the diffusion of the new technological developments and an increasing understanding of the possibilities of the new high-tech tools. For example, marketing research professionals will be more likely than not to want to produce empirical data electronically because it will be easier to produce and diffuse it, and it will be easier to merge the data with other data from other electronic sources.

Summary and Conclusions

Computer-assisted interviewing is asking questions and recording answers in a computer-based environment. It is increasing in importance as a viable high-tech marketing research tool. It can be used in personal, telephone, and mail research designs. The technology can also be used in observation and experimental studies. Its possibilities lie in the imaginations of marketing research professionals.

Computer-assisted interviewing has been possible for 25 years. It was made possible by the introduction of interactive remote terminals and interactive programming. IBM introduced its interactive remote terminal in 1962, just two years after introducing the 1400 Series mainframe computer. Since then, other technological developments, such as WATS, interactive CRT systems, and personal computers, have supported the development of computer-assisted interviewing.

Computer-assisted interviewing offers some basic advantages over other forms of research design. These advantages include bias reduction, error control, and data integration. Specific advantages include randomizing response choices, customizing questionnaires, and producing instantaneous data analysis.

On the other hand, there are some problems associated with computer-assisted interviewing. These problems include hardware costs, debugging time, and respondents without typing skills.

Four factors indicate continued growth of computer interviewing. These factors are technological developments, social trends, computer-assisted interviewing linkages, and a change in marketing research orientation. Technological developments include voice response systems, finger-touch monitors, and video phones. Social trends include the spread of personal computers, the growth of cable and satellite television, and the difficulty of door-to-door research. Computer interviewing linkages include interlocking databases, computer-based models, and automated offices. The marketing research orientation of marketing research professionals will change because it will be influenced by the diffusion and acceptance of the new technological developments, and by an increasing understanding of the possibilities of the new high-technology.

References

Aaker, David A. and George S. Day (1986), Marketing Research, 3rd Ed., New York: John Wiley & Sons, Inc.

Billy, Stephen M. (1981), "Go Slow, Be Wary When Considering Switching to Computer-Assisted Interviewing System," Marketing News, (November 26), 2.

Boyd, Harper W., Ralph Westfall, and Stanley F. Stasch (1985), Marketing Research, 6th Ed., Homewood, Illinois: Richard D. Irwin, Inc.

Churchill, Gilbert A. Jr. (1987), Marketing Research, 4th Ed., Chicago: The Dryden Press.

Dillon, William R., Thomas J. Madden, and Neil H. Firtle (1987), Marketing Research in a Marketing Environment, St. Louis: Times Mirror/Mosby College Publishing.

Groves, Robert M. and Nancy A. Mathiorvets (1984), "Computer Assisted Telephone Interviewing: Effects on Interviewers and Respondents," Public Opinion Quarterly, 48 (Spring), 356-359.

Honomichl, Jack J. (1984), "Survey Results Positive," Advertising Age, 55 (November 1), 23.

Parasuraman, A. (1986), Marketing Research, Reading, Massachusetts: Addison-Wesley Publishing Company.

Tull, Donald S. and Del I. Hawkins (1987), Marketing Research, 4th Ed., New York: Macmillan Publishing Company.

Zikmund, William G. (1986), Exploring Marketing Research, 2nd Ed., Chicago: The Dryden Press.

LONG SELF-ADMINISTERED QUESTIONNAIRES

by
Lesley Bahner
POPULUS, Inc.
Greenwich, CT

In thinking about my topic, long self-administered questionnaires, I asked myself: What is a "long" questionnaire? Given my background in qualitative research, a 30-minute, individual, in-depth interview is short, whereas a three-hour focus group interview is long. For telephone interviewing ten minutes is an ideal length, but 20 minutes is too long. As for computer-assisted, self-administered interviews, I did not know what constituted a long questionnaire. I did have ideas, however, about what makes a "too-long" questionnaire.

An interview is too long if it takes longer than expected to complete. In this case, a respondent finishes an interview upset and angry because you said it would take 20 minutes, but it took him 30. There is evidence that computer-assisted interviewing is perceived to take less time than it does; this perception works to the researcher's advantage, but we should not misuse it.

A "too-long" interview is more than a matter of the time it takes. A "too-long" interview is one which bores respondents. They may say: "I'm bored," "It was too long," "It had too many questions," or "It was too repetitious."

Another indication of a "too-long" interview is when people have difficulty answering the questions. Often this is difficult to determine, because people are unlikely to admit they had difficulty. Instead, they hide their discomfort and embarrassment by saying it was too long or that they were bored. The expression on their faces or body language, however, tells that something was wrong. During the pretest of the questionnaire, spend time talking to these people to find out what was difficult, what they did not understand, or did not know how to do.

One advantage of computer-assisted interviewing is that respondents often regard it as easier and more interesting than paper-and-pencil questionnaires. Yet, there are things we can do to enhance respondent interest, involvement, and ease of completing an interview.

Based on our experience with computer-assisted, self-administered questionnaires, we've come up with five principles for constructing a long questionnaire:

First, people like to tell about themselves. In the interview, give people the opportunity to share information about themselves. Fortunately, this is often the same demographic and lifestyle information we collect.

Second, people want to be entertained. Even in listening to conference speakers, it's more enjoyable if there's entertainment value. The same applies to filling out a questionnaire.

Third, people do not read. This does not mean that they are illiterate, but rather that they do not take the time to read. This applies to those seated before a monitor. People do not like to read screens full of information. There are enhancements you can incorporate into the questionnaire to gain respondent attention.

Fourth, people want respect. They want to feel important and valuable. They don't like dumb questions, questions about unimportant items, or spending time trying to decide what to do. Keep this in mind when writing a questionnaire.

Finally, people don't like to feel judged. Fortunately, computer-assisted, self-administered interviewing minimizes that feeling, since no interviewer is present to raise an eyebrow when the respondent answers. No matter how good an interviewer, a respondent can still perceive that he is being judged. Even with paper-and-pencil interviews, people feel like they're back in school, handing in a test to the teacher. Likewise, in focus groups people often feel I'm testing them when I ask a question. Computer-assisted, self-administered questionnaires minimize feelings of judgment and testing.

This leads to six guidelines for questionnaire administration:

First, maintain a consistency of style while writing a questionnaire. This minimizes confusion, makes the task easier, and takes less time. For example, decide whether to left-justify your questions or center the text. Choose between a 40-column or 80-column format for the entire interview. Explain how to pass from one question to another with similar text and place that text in a similar location in each frame. Also, be consistent about placing the questions on the screen. Such consistency avoids confusion. If you habituate respondents to certain patterns, they will understand each question more quickly and the time to complete the interview will be decreased.

Second, use color. This is very important. Not only is it more interesting and entertaining for the respondent to see an interview in color, but highlighting key parts of questions helps comprehension. Remember, people do not like to read. Making the important parts stand out on the screen catches their eye. For instance, on a set of attribute rating questions, highlighting the attribute (which changes for each question) helps the respondent move rapidly through the list.

Third, pretesting the interview is very important for identifying problems. Take the time to talk to respondents. Many times when you're pretesting a questionnaire, a respondent will complete an interview and say, "Oh, it was great," "It was fun," or "I had a good time." That's all well and good, but you still have to probe the person. Ask him, "What did you think about this question?" Or say, "I was concerned about the scale; did you understand it right away?" Probe them because people forget details that caused them to pause, or they may not admit problems. Take them back through the questionnaire and review the different types of questions with them. When interviewing children, make certain they understand the words. Show them the list of words again and ask, "Did you understand all of these words?" "Are these words your friends might not understand?" Do this in an unchallenging way because people are reluctant to admit that they do not understand. It is up to you to take the time to learn what works and what does not.

The fourth guideline is to use a conversational tone within the questionnaire. This relaxes a respondent. Face-to-face conversations with people make them comfortable and they give high quality information. For this reason, make the effort to "converse" with respondents when constructing a Ci2 System questionnaire.

We do this by naming our program MAX. At the beginning of each interview, MAX introduces himself and speaks in the first person. He explains the subject matter and outlines the task. He also asks the respondent to type in his or her first name, involving the person in the process. Just to show that MAX is paying attention, we restore the person's name in the gender question: "Although I am a smart computer, I still need some help when it comes to sex. So [NAME], are you male or female?" We usually get a chuckle from respondents at this point.

This type of conversation makes the respondent feel involved and important. Furthermore, you can establish a conversational informality in computer-assisted, self-administered interviewing that is hard to achieve in paper-and-pencil interviews, and is especially hard to control in personal interviews. An interviewer reading words rarely sounds casual and personal. But a person reading to himself can interject that element.

We also start with very easy questions like the warm-up of a focus group discussion. While the gender question sounds informal in tone, we use it to clearly explain how the respondent should respond. We instruct them: "Please press the number key of your response: one for male, two for female." These easy questions with explicit instructions allow respondents to become familiar with the computer without becoming intimidated. These introductory questions also give the researcher the opportunity to collect demographic information needed for branching or grouping respondents for subsequent questions.

The fifth practice in constructing a long questionnaire is to include segues, or transitions, between the tasks rather than jumping immediately to the next. If respondents have been thinking about one item for several minutes, let them pause for a moment before asking them about a different item. You can inject humor and visual variety into the questionnaire to relax and entertain them. Make it a short break, only a moment, to pace the respondent. Including one or two such breathers during a long questionnaire makes the questionnaire seem shorter.

Segues provide the opportunity to introduce the next section to respondents, orient them to the type of question, and make them feel comfortable with the upcoming task. If you have 25 lines of introductory text, split it among three screens rather than placing it in one. Again, people do not like to read. Giving them instructions in small segments insures they are more likely to read and absorb them. It is very useful to introduce the scale for an upcoming set of questions. Familiarizing people with the scale before they see the questions helps them focus on the responses rather than trying to understand the scale.

The last guideline is to vary the task. Even though you could use the same type of question to collect different kinds of information, modulate the questionnaire with different question types to reduce boredom and increase interest. Various question types include "pick all that apply," "pick a certain number that apply," and "pick the brand best/least described by the statement." You can also use a staged sort for a long list of items that cannot fit on the screen at one time. By sorting items in stages, respondents progressively narrow down attributes or characteristics that best describe a brand.

Figure 1 shows an example of a "pick a certain number that apply" question type. In this question, the respondent is asked to select those items that describe drinkers of Budweiser beer. "Budweiser" should be in a different color than the rest of the text. Using the Ci2 System, each item is removed from the screen when selected and cannot be selected again.

Figure 1

Choose 4 items to describe
drinkers of

BUDWEISER

The 1st item is

1 Banker
2 Doctor
3 Secretary
4 Auto mechanic
5 Artist
6 High school teacher
7 Rock musician
8 Nurse
9 Librarian
A Chairman of the board
B Social worker
C Army officer
D Kindergarten teacher
E Janitor
F Farmer
G Construction worker
H Hairdresser
I Corporate lawyer

An example of a "most/least" question appears in Figure 2. "Always on a diet" is the first attribute. This line will change with each question, so we make it a distinct color. "Most" and "least" are capitalized and/or presented in different colors.

Figure 2

Thinking of the kind of person who
is always on a diet
which of these four soft drinks
is the person MOST likely to drink?

1 DIET COKE
2 7-UP
3 DR. PEPPER
4 PEPSI

Thinking of the kind of person who
is always on a diet
which of these four soft drinks
is the person LEAST likely to drink?

1 DIET COKE
2 7-UP
3 DR. PEPPER
4 PEPSI

A "pick all that apply" question is shown in Figure 3. The respondent chooses one at a time, as many as apply, and then presses the letter "I" when no more apply.

Figure 3

Which of the following would you use to describe

BECK'S BEER?

Pick all that apply.

1 Is imported from Europe
2 Has a light, smooth taste
3 Goes well with food
4 Has reduced calories
5 Has a rich, full-bodied taste
6 Is available at bars and restaurants
7 Is inexpensive
8 Comes in bottles
9 Is enjoyed by my friends
A Is available where I shop
B Is an American beer
C Can be offered to anyone
D Has a highly distinctive taste
E Lets me drink a lot of it
F Is a Canadian beer
G Comes in cans
H Is a sipping beer
I PRESS WHEN NO MORE APPLY

Figure 4 shows the first list of attributes a respondent would see in the first part of a staged sort. There are 19 items on this list for the respondent to narrow down. In this example, the person is asked to pick eight items that best describe drinkers of Diet Coke. The eight choices are then restored in the next question and the person is asked to pick four. In the next question, these four choices are restored and the respondent is instructed to pick the one that best describes drinkers of the brand.

Figure 4

From the list below, please pick
8 items you think most describe

DIET COKE

The 1st item is

1 Dependable
2 Snobbish
3 Ambitious
4 Flashy
5 Confident
6 Wishy-washy
7 Idealistic
8 Laid-back
9 Impulsive
A Elegant
B Practical
C Satisfied with life
D Sexy
E Energetic
F Nervous
H Ultramodern
I Likes to try new things
J Gentle
K Bright

The first frame of a question using an analog scale appears in Figure 5. We explain the analog scale in the previous frame, then give respondents a chance to practice moving the marker to indicate their answer. We ask: "How completely, or not at all, does good value for the money describe each of the cars below?" After the respondent gives his answer for the Chevrolet Camaro, Toyota Corolla appears on the screen to rate. In this particular analog scale, respondents see their previous ratings as they give their current rating.

Figure 5

Using the arrow keys <- and ->
move the box along the scale to
indicate how completely or not at all

GOOD VALUE FOR THE MONEY

describes each of the cars below.

DESCRIBES NOT AT ALL   DESCRIBES COMPLETELY

CHEVROLET CAMARO

TOYOTA COROLLA

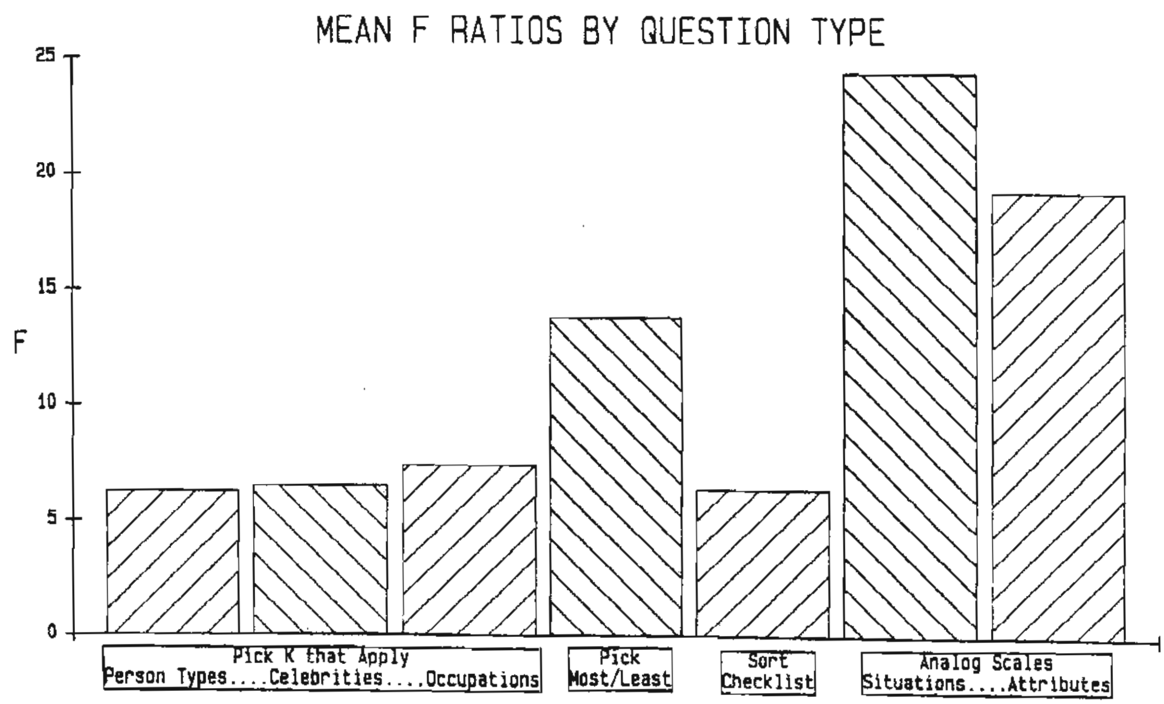
FORD TAURUS

YUGO

The analog scale has been somewhat controversial during this conference. I've heard disparaging comments about it, yet respondents like it. In fact, that's the question-type they remember most. They describe it as "a lot of fun." Some critics are concerned about the length of time it takes to answer an analog scale question, since the respondent has to move the marker right or left. When we pretest a Ci2 System questionnaire in the office, we get bored moving the cursor back and forth. But most consumers are unfamiliar with computers and do not perceive it as being too slow. Moreover, the creator of the questionnaire has control over the length of the line, which does not need to extend across the entire screen. Also, you can tell respondents that holding down the arrow key moves the cursor faster.

More concrete evidence that the analog scale works comes from calculations of F ratios for the question types just described (Figure 6). We conducted 1,200 Ci2 System interviews with each person rating four brands on 160 items. In this particular case, a ratio greater than two was significant. Of the four question types, the analog scale differentiated more than the other types of questions.

Figure 6



Other researchers question whether at the end of a long interview bored respondents "press buttons" to finish. That may be true in some cases, but I doubt it occurred in this study (Figure 6). One of our analog scale questions appeared early in the interview, whereas others came close to the end. Even at the end of a long questionnaire, analog scales hold respondent interest.

Finally, there are several other concerns about administering long questionnaires for which I have only partial answers. One concerns the time required to complete an interview. When you get data from an interview that should have taken 40 minutes, you may find that some people took 20 minutes to complete it and others took 60 minutes. Ask yourself: "In which cases do I have good or poor quality data?" The second part of that question is: "Is it a respondent problem or a questionnaire problem?"

To answer these questions, consider the following: Look for the respondent who is just pushing buttons. You may find a random pattern of response in the data, or an obvious pattern, such as 1-2-3-4-5-6-5-4-3-2-1 for six-point scale questions. Then look for patterns of interview times by facility. Once we got back questionnaires and found one facility had a series of consecutive 12-minute interviews on one diskette. We called the facility and tracked the problem to a dishonest staff member.

Also look for patterns by demographic group. For example, when interviewing a wide age group, eight to ten year olds may have difficulty understanding certain questions that adults understand readily. If the data do not make sense by demographic groups, the problem may be respondent comprehension, not data analysis.

Too long a questionnaire? It won't happen if you make it an interesting conversation with the respondent.

POLITICAL POLLING

by
Brent Stahl
MORI Research
Minneapolis, MN

I am here to discuss the practical aspects of a series of election day polls MORI Research conducted using the Ci2 System. These studies are interesting for two reasons: They illustrate what is possible when ambitious goals must co-exist with modest budgets, and they required probably the largest number of terminals used to date for a Ci2 System project. I will conclude with some thoughts about the role of computer-assisted interviewing in political research in general.

BACKGROUND

In the past few years, print and electronic media organizations have frequently combined forces to sponsor public opinion surveys. In the Twin Cities of Minneapolis and St. Paul, the St. Paul Pioneer Press and Dispatch, WCCO Radio, and WCCO Television have produced the "Northstar Poll."

A continuing problem for these organizations has been how to cover the primary and general election voting. In close elections there was little to report except anecdotes until most of the returns were in, which often could be the next morning.

In the 1978 primary, the St. Paul Pioneer Press and Dispatch embarrassed itself by declaring in a banner headline that the U.S. Senate candidate leading comfortably at midnight had won the election. Unfortunately, that candidate lost by 3000 votes because of extraordinarily one-sided voting in the remote areas of northern Minnesota.

In later elections the news organizations used a modeling procedure in which election returns from selected precincts were analyzed on the basis of known demographic characteristics and previous voting patterns. This procedure had some advantages, but it was still limited by the need for raw vote totals.

In planning the 1986 election coverage, managers at the three news organizations decided they needed some type of election day survey of voters. We at MORI Research were commissioned to conduct the research, as we have considerable experience in media research and election surveys.

We were given two goals: First, to provide information that would allow the radio and television stations to project winners by 8 p.m. (when polling places close), or at least to caution them that they should not project a winner. Second, we would provide data for analyzing the governor's election in terms of the effects of issues, demographics, and perceived characteristics of the candidates. This was important for all three sponsors but especially for the St. Paul Pioneer Press and Dispatch, since most of its readers would know who won the election by the time they read the newspaper.

PLANNING THE STUDIES

The first issue to resolve was what kind of survey to do. One way, which MORI personnel had done before, would be an "exit poll" similar to those done by the TV networks.

This procedure was developed initially by Warren Mitofsky of CBS News. It requires selecting a group of representative precincts, ranging from strongly Democratic to strongly Republican. Interviewers systematically select voters as they leave polling places and ask them to fill out a brief (6-8 questions) instrument. The interviewers periodically call in the answers, which are added to the database.

Exit polling works well for choosing winners and estimating the importance of some issues and candidate characteristics. However, it has the disadvantage of being quite expensive and logistically complicated to manage. Also, length requirements limit the range of questions that can be asked, which would be a problem for our newspaper client. For those reasons, we settled on interviewing voters by telephone.

One may wonder about the prudence of doing a telephone poll of voters on election day. We at MORI Research have done 10 such surveys with excellent results in Minnesota and Chicago areas with very different political cultures. These include elections ranging from mayor to U.S. President. These studies require very careful attention to sampling procedures and question ordering.

They also require examining local conditions. Election day telephone polls are feasible in Minnesota for three reasons:

- The voting rate is quite high - often the highest in the country. Turnouts of 60 percent for governor's elections and 75 percent for presidential elections are not unusual.

- Almost all households are reachable by telephone. According to the Public Utilities Commission, 96% to 98% of Minnesota households have telephones. Those without telephones are probably not likely to vote.

- Minnesotans, like other Midwesterners, are fairly tolerant of telephone interviewers, compared to residents of some other areas of the country.

WHY USE COMPUTERS

After deciding to go with telephone interviews, the next issue to resolve was whether to use computer-assisted interviewing or conventional paper-and-pencil questionnaires. The latter was not feasible because of our need for both speed (the 8 p.m. deadline) and depth in the questionnaire, which included 30 questions. If speed were the only requirement, we could have used a short paper-and-pencil questionnaire and had the data keypunched as the interviews were completed.

The ability to field computer-assisted surveys is a selling tool for us. In this study, the Ci2 System questionnaire helped us sell the study to an audience that normally doesn't need persuasion - the interviewers.

We determined it would take 75 interviewing stations to complete 900-1000 interviews on primary election day. In Minnesota the ratio of Democrats to Republicans runs 2:1. We were not expecting a close race among the Republicans, so about 300 interviews from that population seemed adequate. We wanted at least 600 Democratic interviews, since that race would be closer.

Seventy-five interviewers, even for one day, are expensive, so the sponsors decided to use their own personnel for this task. This included reporters, editors and producers who were not working on election-day stories.

Many of us at MORI Research have worked in media companies, so we knew having these people as interviewers could pose problems. Indeed, many of those "volunteered" initially failed to appreciate what an honor it was to be among our interviewers!

This was especially true for the editors and reporters from the St. Paul Pioneer Press and Dispatch. A contributing factor was that contract negotiations between the newsroom employees' union and company management were not going well at the time. It was quite possible that the union would walk out the weekend before the primary election.

Newspeople tend to be intelligent and independent; they don't like to be told what to do; and they don't like their routines to be disturbed. We came in as outsiders to disrupt their routines and instruct them how to conduct a proper interview.

Our assigned interviewers were busy professionals, and we had just four hours to train each group from the three sponsoring organizations. We had to spend much more time than usual explaining the value of survey research and "selling" this particular survey to the interviewers.

Our Field Director is very good at this type of instruction, but the interviewers were further convinced when they sat in front of the terminals and saw how easy it was. All they had to do was read a telephone number sample sheet, ask questions in a neutral voice, and press a key. While still not thrilled to be on the project, they were clearly more accepting after their exposure to the Ci2 System interview.

NUTS AND BOLTS

A not insignificant task was to find 75 computers for our interviewing stations as well as a place to put them. The Minneapolis rental firm we used had never had such a large order, and they had to rent machines to fill our order. Our clients, ever mindful of out-of-pocket costs, traded out much of the computer cost in advertising time.

For our base of operations we needed adequate space and power for the computers and a place for the television people to do remote location telecasts. We considered several locations, including a recently closed hotel, but settled on a storefront in downtown Minneapolis that had once been a Digital Equipment Corporation retail store.

The Minnesota Primary was held on September 9, 1986. Our interviewing went well, but we had to deal with various practical problems during the course of the day. This is true of any survey, of course, but the brief time we had for interviewing increased our urgency in working through the problems.

TERMINALS

We had our first look at the rented computers on election day and were relieved that they were delivered! Since we were expecting IBM PC's with both color and monochrome monitors, the interview was programmed to display color. However, about an hour before interviewing, we discovered that 20 of the PC's were equipped with a color card and an off-brand monochrome monitor that displayed colors as shades of green. This would not be a problem on a high-resolution monitor, such as those used by Compaq computers, but the interview was unreadable on these monitors.

I quickly removed the color statements from the original program disk, but there wasn't time to copy the revised program onto each of the 225 field disks. As a result, we kept a separate batch of disks for the 20 problematic terminals. This was inconvenient but manageable.

PRACTICE INTERVIEWS

The Ci2 System questionnaire disks have a "practice" mode in which the interviewer enters a zero as the interview number. The questions appear as usual, but the answers are not written to the data files.

Despite repeated instructions, we found that some interviewers were doing practice interviews in the "real" mode. This forced us to check with each interviewer for practice interview numbers and to purge those interviews later, using the data management utility of our tab software.

CUMULATING DISKS

We needed to take stock of our results at three times - 5 p.m. to check for problems; 7 p.m. to prepare for the 8 p.m. newscasts; and 9 p.m. at the end of interviewing. The Ci2 System CUM program transfers interviews from individual disks to a master database easily and fairly quickly. However, we had to run through 75 disks three times and this procedure grew old quickly, especially as our deadlines were approaching.

WATCHING THE SAMPLE

We used random-digit dialing to reach households and alternated between asking for the youngest male or youngest female who had voted. The questionnaire branched to Republican or Democratic questions depending on the election. Our sample duplicated the 70% - 30% split between Democratic and Republican turnout.

Part of our sampling strategy assumed that the voting turnout would vary in different areas of the state because of the candidates involved. (Assuming equal voting rates is often a source of inaccuracies in pre-election surveys.) We, therefore, stratified the sample roughly by the eight Minnesota Congressional districts. Since we were covering both Republican and Democratic elections, we had 16 district quotas to fill.

Our task was considerably complicated by the need to maintain as geographically "balanced" a sample as possible during the course of the day. We wanted to look at our results at 5 p.m. and then again after 7 p.m. in order to make the 8 p.m. newscasts. Another consideration of interviewing on computers was the potential problem of a power outage. It might have become necessary to go with a partial sample.

These considerations made us continually monitor our sample, a difficult task because of number of interviewing stations and quotas involved. At times the interviews in one or more zones came much faster than in other zones, and we had to switch telephone sampling sheets in order to maintain the balance.

THE TEDIUM OF DIALING

The only complaint we had from interviewers was that it was boring to dial manually from a list of random telephone numbers. As is always the case with this procedure, we had many non-working numbers.

We mentioned this fact of life while training the reporters, editors and producers serving as interviewers. We did not dwell on it, however, as it detracts from the glamour of telephone interviewing.

The Ci2 CATI System has utilities that easily handle the problems of cumulating data from disks, quota control and telephone dialing. However, the budget for this study would not support the extra expense of the CATI System.

OPEN-ENDED QUESTIONS

As part of the interview we asked respondents to say in their own words why they voted for either candidate. We wanted verbatim responses to serve as anecdotal material for news stories. "Keywords" would not be helpful.

Some of our interviewers were not used to taking dictation or could not type well, so their answers often were not interesting or useful.

A number of the interviewers, especially from the newspaper, were used to writing on terminals. We thought it would be easy for them to type in the open-ended responses in the Ci2 System. This was not the case. While reporters are used to typing while interviewing, they often use their own shorthand or type in a fairly haphazard manner, then edit their work immediately. They could not edit their Ci2 System answers, however, so many of the results were not usable.

As a result of this experience in the primary election, we changed the procedure in the general election survey to have only experienced reporters ask open-ended questions and record the answers by hand.

THE RESULTS

In election day surveys, the only numbers that anyone remembers are the vote totals. As the following tables show, we were quite close to the Democratic result and close enough to the Republican totals. We probably would not have surveyed the latter primary except that Republicans had criticized the local media for ignoring their election.

DEMOCRATIC PRIMARY (N=621)

	ACTUAL	POLL (7 P.M.)	POLL (9 P.M.)
Perpich	57%	58%	57%
Latimer	41%	37%	38%

REPUBLICAN PRIMARY (N=266)

	ACTUAL	POLL (7 P.M.)	POLL (9 P.M.)
Ludeman	77%	68%	69%
Lindau	16%	16%	14%

We also obtained considerable information on the strengths and weaknesses of the candidates and why the Democratic election in particular was not closer.

AN UNEXPECTED SCOOP

Most of the news organizations in the state had contracted with the Associated Press (A.P.) to provide them with vote totals using a rather elaborate computerized reporting system. Minnesota is a geographically large state with more than 4,000 precincts. With two parties and many offices involved, counting and reporting the votes is not an easy task.

Disaster struck early in the evening when the A.P.'s mainframe computer went down for the count. They were reduced to hand-tabulating results from 4,000 precincts, and when the local newspapers went to press after midnight, just 20% of the vote totals were available. Our survey provided the only good information on election night, and other local media had to use our results (properly credited, of course). Our clients, naturally, were quite pleased with this development.

THE GENERAL ELECTION

We expected that the November General Election study would be much easier to conduct for several reasons:

- There would likely be a much higher voting rate, perhaps as high as 50%

- There would be one general election rather than two primaries, requiring one database

- Our interviewers had the experience of the first survey

The interviewing was relatively uneventful, and we again came quite close to the actual margin:

GENERAL ELECTION (N=1003)

	Actual	Poll
Perpich	57%	56%
Ludeman	42%	42%

COMPETITIVE RESPONSE

We were hoping, of course, that the A.P.'s computer would fail again, but it worked well.

Probably in response to our work, one of the competing Minneapolis television stations conducted its own general election survey "with a difference." The survey they reported on election night was conducted in the 10 days before the election. Since the political environment in Minnesota did not change appreciably in that period, their results had a degree of face validity. However, this was not a project that I would have wanted to do.

COMPUTERS AND POLITICAL POLLING

Political polls for media organizations such as those described here are fun to do, even though you live close to the edge at times. The most interesting political work to me is not for public consumption, but rather for candidates and incumbent politicians. These surveys are typically designed to produce much more detailed information than is possible in media polls.

Whatever the type of political survey, microcomputer-based interviewing programs do not allow one to ask questions or perform analyses that cannot be done otherwise. They provide advantages in speed and costs. Microcomputing interviewing is faster, if not cheaper than paper-and-pencil procedures. It is cheaper, if not faster than mainframe or minicomputer based interviewing systems. These advantages are significant.

Speed is increasingly important in political research, especially as television has become the primary medium for political messages of all kinds. In an age when the political situation routinely changes overnight, politicians (and media organizations who follow them) have increasingly demanded the capability to monitor these changes. Computer-assisted interviewing packages allow this need for timely, sophisticated research to be met with less effort and expense than has been possible in the past.

A CLIENT'S EXPERIENCE WITH COMPUTER INTERVIEWING

by
David Griscavage
Pepsico, Inc.
Valhalla, NY

Computer interviewing represents an important technology at PepsiCo. Prior to our acquisition of the Ci2 System, we investigated most of the computer-assisted interviewing software on the market. Our objectives in using computer interviewing are:

-To Decrease Turnaround Time for Research

-To Improve Testing Procedures

-To Decrease Interview Time

-To Heighten Interviewer and Respondent Interest

The decrease in turnaround time is to be accomplished during the fielding and data acquisition phase. Computerized interviewing eliminates the need for keypunching and data cleanup, which are frequent bottlenecks for quick turnaround.

Specifying exact testing procedures in an unambiguous manner has always been a daunting task in questionnaire development. For this reason, novel methodology has always been difficult to implement with paper questionnaires.

In contrast, computer-assisted interviewing (CAI) provides explicit control over interviewing activities, including skip patterns and prompts thanking the respondent at the end of an interview. Improving testing procedures is potentially one of CAI's strongest advantages when compared with paper-and-pencil interviewing. CAI also opens up the potential for self administered interviews, leading to lower costs and no interviewer bias.

Long interview times are one of the more unpleasant aspects of the research experience. CAI may be an effective tool in helping to reduce the real (and perceived) length of interviews. The automatic sequence of CAI undoubtedly will lower interview length compared to identically structured paper interviews.

PCs are currently available in just 15% of U.S. households. For many interviewers and respondents, the machine represents a "high-tech" experience. PCs allow the use of handsome graphic and color questionnaire presentations that pique interviewer and respondent interest. For the interviewer, using a PC represents technical mastery of a state-of-the-art research tool and avoids tedious paper work.

Roles of Client/Supplier/Field Service

Since we at Pepsico write most of our questionnaires, we find that suppliers and field services are invaluable in providing a check on questionnaire materials. All too often under deadline, things tend to slip through the cracks. Our suppliers/field services have provided us with important feedback on questionnaires, leading to better questionnaire development.

Suppliers/field services are in an excellent position to procure the hardware for computer interviewing. In most cases, they have developed good working relationships with local computer shops. Our role is to specify minimum requirements for our job.

It is difficult for us to provide suggestions such as how many computers to rent or the maximum number of field diskettes that are necessary. Input on these issues helps us to fully understand the field operation and, more importantly, justify the costs internally.

Field services must recognize that a Ci2 interviewer must have additional training. We look to field services to conduct training among the individuals who work on our interviews. Additionally, we would like to see an industry-wide training standard developed. We fear that some field services may hang out an "I Do Ci2" shingle without training their personnel.

Computer interviewing need not be more expensive than conventional methods. While initial equipment costs are somewhat high and training/development of personnel can be expensive, I believe it is unwise to recover these costs over a short period of time. The risk of alienating potential clients carries a cost which is not easily overcome. We urge newcomers to computer interviewing to be judicious in how they cost computer studies.

Comparisons of Computer Interviewing vs. Paper-and-Pencil

We have found that there is a longer lead time needed for Ci2 System questionnaire development compared to paper-and-pencil. This is because there is no "word processing" analog procedure available for Ci2 System questionnaire development. With paper questionnaires, "boiler plate" documents are available and can be easily revised. While Ci2 System questionnaires can be changed, their modification is not easy.

As is the case with paper questionnaires, Ci2 System questionnaires need a review process prior to fielding. This review must be done by someone with Ci2 expertise, since all possible program routes and skip patterns need to be evaluated. This process can be time consuming since each skip pattern must be followed to interview termination points.

IBM PCs are the hardware of choice. Field sites must obtain PCs with graphic boards (and color monitors when necessary). IBM clones can present problems especially when COMMAND.COM file requirements are different. We specify IBM machines on most jobs.

With the advent of computer interviewing, field instruction must include "boot up" and simple DOS instructions so that inexperienced field staff can operate the Ci2 System. We only use the Ci2 System with experienced field services so that computer operation basics are not an issue. We do have instructions on the use and dispositions of field disks that are particular to our needs and include these in the field instructions.

Our experience with hardware rental is that the quality of the machines available is quite variable. We have obtained machines in which the A drive did not work and monitors were inoperative. Field services need to deal with rental agencies that rent machines as part of their business (i.e., Computerland). We would not recommend shipping equipment - it is very costly and potentially harmful to the hardware.

One must decide at the outset how to differentiate interviewer instructions from the interview text. Prompts for interviewer activities (obtaining products, displaying cards) can be done easily on the Ci2 System through the use of separate screens, different colors, or 40/80 column text. Instructions in 80-column format and text in 40-column format is my favorite method for denoting the difference to the interviewer.

Costs for doing computerized interviewing are highly variable. Experienced field services tend to provide the best value overall. Inexperienced field services do not bid the computer work well. Their lack of experience about rental costs and other charges related to computer interviewing often are reflected in extremely high bids. Often it is helpful to obtain a paper-and-pencil analog bid to determine the differential for computer interviewing. When comparing bids, factor in the cost for keypunching paper questionnaires to obtain a true measure of comparison.

Issues During Fielding

Currently, there is no method for maintaining controls for screening or product quotas using computer interviewing software. Thus, a portion of the research must still be implemented using paper materials.

Everyone remembers his first experience with a computer. For some interviewers and many respondents, our interview represents their initial experience with the machine. For the most part, the Ci2 System makes the experience painless. But for some, the computer is still threatening. We suggest using practice interviews on all new jobs, especially when introducing new interviewers to the PC. Some respondents may never feel completely at ease; for them, interviewer-conducted interviews are the only solution.

On the other hand, a young computer wizard may try tricks with your program. Keep this in mind when programming and check how your questionnaire handles out-of-range answers.

Having all interviewers run through the Ci2 System interview is a must. Interviewers should be encouraged to thoroughly explore the entire interview including all skip patterns and out-of-range responses. This activity has two purposes: 1) it gives the interviewer an understanding of the interview and its goals; and 2) it serves as a final check on the programming.

The key individual responsible for the proper execution of a computer interview is the supervisor. We strive to maintain close communications with field supervisors, since we're aware of their importance to successful job completion. With computer interviewing comes additional supervisor responsibility. It is important to develop a rapport with the field supervisor to keep informed of the progress and status of the project.

We currently conduct only interview-administered interviews that sometimes contain a self-administered segment. Central-location taste tests do not easily lend themselves to a self-administered format because of the logistics of product presentations and the monitoring of product consumption. We've found that the periods when products are being readied for presentation are excellent times to conduct self-administered sections containing demographic and product-use information. These breaks between product presentations are good times to keep the respondent busy in a desirable activity.

An issue with obvious ramifications is liquids and computers. Both client and field service need to think through the position of the hardware, especially the keyboard, with respect to the products being tested. Cups of soda are inevitably spilled. They tend to be spilled in a manner that causes most harm. Do not forget to include your diskette as an item liable to ruin from spilled products.

One potential drawback of computer interviewing is also one of its major advantages: the ease of data entry. It can be so easy that an interviewer may not carefully key-in the data, leading to a substantially large number of "mis-strokes" in the data. While we have not yet assessed the depth of this possible problem, several preventive steps might be used. Highlighting and recalling answers are ways to make the interviewer aware of mis-strokes. The ability to x-back also helps interviewers to correct mistakes. Nothing can be more effective than interviewers who take care with their work.

Data analysis at PepsiCo is generally conducted using software such as SAS run on a VAX machine. To use a mainframe it is necessary to upload the data set from a diskette to a mainframe file area. Several software packages (e.g., Smarten) are available for this purpose. However, these packages assume that your PC is connected to the mainframe via a hard wire or modem. Your local technical support person can be helpful in solving any problems with file transfer.

Very little cleaning of Ci2 System data sets is generally necessary. We do a validity check on the screening data with demographic/product use information that we ask during the interview. The demographic/product use data should correspond with the screening data for that respondent. In the majority of cases, we have found good correspondence, meaning that the field services are doing a good job. If a large amount of noncorrespondence was found, we would attempt to determine its cause with the field service.

We have only had one problem with temporary data loss. That problem was solved by the staff at Sawtooth Software. Norton Utilities is a software package that might help recover accidentally erased data.

TELEPHONE INTERVIEWING

by
Richard Miller
Consumer Pulse, Inc.
Birmingham, MI

I've been asked to discuss traditional telephone research through the use of computers, and what Consumer Pulse has learned so far. Our company comprises mall intercept locations and field offices. We operate in excess of 180 telephones throughout the network of offices for local telephone interviewing. We also operate WATS centers in Cleveland, Detroit, and Denver for conducting interviews nationwide. Our WATS Center in Detroit is equipped with the Ci2 CATI System with ten networked IBM PCs and ten additional stand-alone IBM PCs.

Telephone interviewing is the most popular form of interviewing, with estimates that nearly half of all marketing research is conducted by telephone. There have been two basic ways of conducting telephone interviews - on paper and with the use of computers.

The use of IBM PC computers provides a relatively inexpensive way to collect data. We know that computer interviewing data are the cleanest data money can buy, and we can generate instantaneous tabulations of any size study. We have had a 3,000 sample study in which we started to collapse data from the field disks at 9:00 am and generated hundreds of sheets of tables by 4:00 pm.

Being from Detroit, we have been involved with automotive research. Automotive research presents some of the most interesting marketing problems. In one series of studies we conducted for an automobile manufacturer, the clients were likely candidates for the Ci2 CATI System, but they had to be SOLD. Here's how we discussed Ci2 CATI with them.

They had a need for information about particular features of automobiles. We were told we would survey subjects about radios, wipers, trunk room, and so on. Normally, we would collect 200 to 1000 interviews on a particular issue, depending upon the number of automobiles involved and the extent and importance of the issue.

We suggested that conducting the interviews on computers would allow us the maximum flexibility and quality control in obtaining the interviews. More importantly, we could provide verbatim comments to the engineers, rather than coding the open-ended responses.

Product planners and engineers of automobiles have had an extremely difficult time with open-ended data over the years. They have received coded tabulations which say, for instance, "I like the front-end styling," "I like the rear-end styling," or "I like the interior dashboard." Even detailed coding would provide responses, such as, "The dashboard is laid out well," "The radio is in the correct location," or "I like dual headlights." Coded open-ended responses for product planners did them little or no good.

The product planners were mostly interested in verbatim comments - long, verbose, open-ended responses. With the Ci2 CATI System, Consumer Pulse can now provide verbatim open-ends to product planners economically and quickly. We can sort the open-ends into a variety of subgroups, including the make of car owned, frontview styling, sideview styling, rearview styling, interior styling, dashboard styling, etc. Or we can code the open-ended responses and give subtotals of ten people who say they liked the location of the radio. The verbatim response indicates exactly what respondents said.

We told our automobile client that the verbatim open-ends could also be sorted using key words. In one study, respondents were asked for two or three key words which came to mind when we read them a name. One of the names related to a magician, and typical open-ended responses were The Wizard of Oz, Oz, a magician, and many derivatives of these key words. Utilizing the Ci2 System's open-ends, we were able to sort the open-ends into groups of key words. The client then saw exactly what was said and the context in which it was said. This information was more valuable than a coded closed-ended response, which could potentially group Wizard of Oz and Oz together.

We also discussed the incredibly complex skip patterns with the product planners. For example, automobile advertising recall is one of the more difficult types of studies to administer on paper. With computerized interviewing, the interviewer no longer needs to wonder which question should be asked and which should not. In addition, complex skip patterns, in which whole sections of questionnaires may be asked or not asked based on previous responses, are correctly handled using computerized interviewing. The cross checking and editing to assure that sections of questionnaires have been asked are automatic.

Automobile telephone research also has an incredible number of evaluations. The Ci2 System quickly and easily handles the rotations and randomization of these evaluations. We told the product planners that they no longer had to worry about positional bias with computerized interviewing.

The Ci2 CATI System offers something we have found in no other computer-interviewing system. Some of our clients have some resistance to doing computerized interviewing because they cannot see the computer questionnaire. More importantly, in most WATS centers that utilize mini-computers and terminals, clients have to travel to the data collection site to view a computer interview being administered. Using the Ci2 System we prepare the computer questionnaire, send the client a field disk, and he can use any PC to view his questionnaire and check its correctness. We know of no other computerized system that utilizes this feature and still offers all the advantages of the Ci2 CATI System.

The product planners knew the most important factor in conducting their telephone interviews with the Ci2 CATI System was quality. We told them quality could be achieved with no increase in total project costs compared to paper, as long as we put up the questionnaire, conducted the interviewing, and tabulated the results. Today, a field and tab project is about the same cost on paper as on computer. But more importantly, the quality of the computer-interviewing data is greater than that administered on paper. Our tabulation staff would much prefer to have computer-interviewing data than data that requires keypunching of paper questionnaires.

We sold the product planners on computer interviewing. But there were other logistics and questions that had to be addressed.

For most automobile studies, our clients design the questionnaire, and our spec writers turn it into a computer questionnaire. Standard questionnaires are approximately six to eight pages long and take 10 to 15 minutes to administer. At most, it takes one or two days to create a computer questionnaire using the Ci2 System. We may or may not elect to conduct a study on the CATI System. Out clients in the automotive industry do not always require tally information, and we can easily administer their interviews using stand-alone computers. However, the information from CATI for the tallies is useful for monitoring individual interviewer productivity and incidence factors, and the CATI System handles call backs to respondents perfectly.

Since their questionnaires are straightforward and have already been approved on paper, we usually go directly to conducting the interview with a few pretests. However, before interviewing begins, we send a copy of the field disk to our client for review, and they call us with approval on the computer questionnaire.

Interviewer briefings are much easier using computer systems. The briefing becomes a round-table discussion of an overview of the study with some specifics about the study detail. Most of the briefing occurs in the computer room. Having eight to twelve people sitting at the computers working together conducting practice interviews is efficient and economical. Sometimes two or three interviewers sit at a terminal and our briefing leader instructs the interviewers which buttons to push. All the interviewers are being briefed in the same way.

Interviewer reaction has been favorable. They are required to push a 1, 2 or 3. All our interviewers can easily enter closed-ended answers into the computer. In addition, almost all interviewers today have some typing skills.

Open-ended responses are diamonds within the research. However, the diamonds may be initially rough. There can be typing errors, the letters may be upper and lower case, and the punctuation marks may or may not be correct. Probe marks are somewhat difficult for the interviewer to insert in the computer, and some interviewers are not proficient typists and spellers. As a result, we have to carefully screen computer interviewers. On studies with large numbers of open-ended questions that require probing, we use interviewers with good typing skills. In addition, Consumer Pulse has had to develop procedures to "clean up" and sort the open-ends for final reports.

We have not encountered major problems with the Ci2 System in telephone interviewing situations. In fact, the problems on paper have been greater than on computer. However, much of the reason for the success is attributed to good planning on the part of clients and our personnel, and good execution of the computer interviewing questionnaire by us.

On the other hand, we are often the subcontractor for collecting data using computers in telephone and personal interview research. We have encountered most of our problems in this area, which generally relate to clients who do not use the proper DOS in formatting their field disks. Other clients do not provide instructions within the computer interview. For example, some clients have self administered interviews without instructions on the use of the "X-back" option. Others have used numbers out of range for refused answers in self-administered interviews and expected our interviewers to remember this. One client even set up study numbers in questions at the end of the interview. He expected us to enter the three-digit code in order to end the interview and record it on disk, but he never told us the code.

Skip patterns have been missed when our clients have put up questionnaires. In one study, we conducted 1500 telephone interviews and the key question in the questionnaire was skipped over. All 1500 respondents were recalled to get this key answer. Of course, our client had to explain the problem to his client, and a considerable sum of money was spent to contact the respondents. We were thankful we had not designed the questionnaire or created the field disks.

Returning to our automotive studies, we have conducted 20 in the past year. Our pricing for a computer study is basically dependent on the number of interviews that are completed, but for more than 300 interviews it is definitely more economical to use a computer interview, unless there are a large number of verbatim open-ended responses. With at least this sample size, our cost for conducting computer interviews is essentially the same as for paper. More importantly, our overall turnaround time including tabulation improves by three to four days because of clean data and a verbatim open-ended file. Also, the front end time in converting the questionnaire and creating the field disks is approximately the same as retyping the questionnaire, having it printed, and marking rotation starting points.

Consumer Pulse was one of the test sites for Ci2 CATI System. We have found the Ci2 CATI System to be one of the best CATI Systems on the market today. It does almost everything that I could ask a CATI System to do given its price and flexibility.

In using the Ci2 CATI System, we have found one major problem - the file server must be a dedicated PC AT. Unfortunately, our PC ATs are in high demand. They are fast, users like them, and they tabulate studies very quickly. Our PC ATs are used daily, morning until night. This means that our dedicated file server is not so dedicated. I have had a request to purchase an additional PC AT as the file server, but my fear is the new PC AT will also be utilized day after day.

Respondent quota control is superb using the Ci2 CATI System, particularly for a large study that requires additional computers beyond the networked CATI System. We have used the networked PCs and an additional six or eight PCs in stand-alone mode to collect data for the same study. This requires us to adjust the quotas in Ci2 CATI downward daily. It is, however, a simple solution to handling large studies that use more computers than the network can support.

Although the sample database in Ci2 CATI is an excellent product, we are not utilizing it fully. We have yet to generate a totally random sample, nor have we directly input a computer-generated sample given to us by the client. So far, the interviewers control the sample, not the CATI System. We anticipate that this will change as we grow with the CATI System.

As a data collector, I have some tips on setting up your facility to maximize the use of your computer interviewing system. First, computers generate a great deal of heat, particularly color monitors. Whatever your architect designs for air conditioning, double or triple the volume of air flowing into your telephone rooms.

Second, lighting is a very critical problem in telephone rooms with computers. Overhead lights cause a great deal of glare on the screens. Using back lighting with no overhead lighting affords the least glare on the screens of the computer monitors.

Next, desk space is at a premium at a telephone center. Desks need to be at least 30 inches deep. Keeping the central processing unit on a shelf above the desk is the best location. This frees up desk space, assures coffee will not be spilled on it, and the unit will not be kicked over.

Of course, soundproof materials are extremely important in a telephone center. I must stress the importance of the use of these materials with the clicking of the keys and the beeping of computers. With 15 or 20 computers beeping in a telephone room, interviewers do not know whose it is unless the room is well soundproofed.

Electrical power coming into a telephone center can be a problem. Many offices have single outlets and the power capacity is designed for operating calculators, coffee machines, and typewriters. In a telephone center with 20 or 30 PCs with color monitors, the power requirements may be greater than your office provides. It is extremely important to assure that you have an adequate amount of power, but also a great number of plugs on separate circuits.

Unlike mall interviewers, or the occasional user of a computer, telephone interviewers look at a computer screen for four to eight hours a day. Color is important in a telephone center. We recommend the use of EGA color monitors for telephone centers. Although these monitors are a bit more expensive, they are easier on the eyes and maintain the quality control that color affords. On the other hand, we have had a recent report of eye strain among our interviewers. This has been attributed to color monitors emitting a large amount of radiation. We are looking for protective screens to put over the monitors to help eliminate eye strain and excess radiation emissions.

Finally, the equipment selection must be a major consideration. As those of us who work with computers know, there are major changes coming for PCs. We still believe that the IBM Disk Operating System (DOS) will be the standard in the entire industry for the next three to five years. However, major changes will appear in DOS, and IBM will still lead the way. We are uncertain of the compatibility of non-IBM computers in the future. As such, we have elected to stay with IBM equipment exclusively. IBM has provided us with excellent technical and hardware support. The same technical support may not be available for non-IBM machines.

Overall, the telephone center at Consumer Pulse has been and will continue to be an extremely exciting venture. With the unlimited potential of Ci2 CATI, we anticipate adding additional machines and locations throughout the United States. With the ease of use of the Ci2 CATI System and its transportability, we could easily have 180 central telephones calling nationwide. Given the increased problems of hiring interviewers in selected labor markets or coordinating nearly 200 persons to show up at a single telephone center, the concept of decentralized WATS with an interactive computer system using Ci2 CATI seems practical for the future.

CHILDREN'S RESEARCH AND OTHER ISSUES RELATED TO THE USE OF COMPUTERIZED INTERVIEWING

by
Ira W. Goodman
J.M.R. Marketing Services, Inc. and
Its Youthful Insights Division
New City, NY

The primary focus of my talk will be on the use of computerized interviewing in research conducted with children. I will also touch on two other issues: 1) The use of computer timing capabilities - or how to use a function created for one purpose in a totally unanticipated fashion; and 2) my view of self-administered computer interviewing, which may differ somewhat from those expressed by my colleagues.

The research efforts of our Youthful Insights Division for the last year-and-a-half have spanned simple concept studies, concept screening tests, and product and advertising testing. The ages of the children tested have ranged from 3 to 12 years. Sometimes the studies involved testing a child and then the mother. Other times, only the child was tested. However, all of our studies with children have been conducted in central locations with the Ci2 System.

You are probably conjuring some wild images of children playing with computers and asking yourself, "But how can they read the questionnaire?" If you have these notions in your mind, it is probably because you believe that one of the key benefits of computerized interviewing is the absence of an interviewer. While self-administration and fast turnaround of results are benefits of the system, our children's research procedures rest on the field execution advantages of the system.

Field execution advantages are all the procedures that a computer interview can control to simplify the life of the interviewer and to assure that the study is carried out just as it was intended. For example, a paper questionnaire with complicated skip patterns can be very confusing even for the best of interviewers. The computerized interview automatically skips to the appropriate questions. This allows the interviewer to remain calm, not feel overwhelmed, and concentrate more on relating to the child. As a result, the child feels more at ease in the testing situation.

Another simple field execution function involves randomization of products, concepts, or commercials. A paper-and-pencil study usually handles this with different versions of the questionnaire. Although this does not represent a major problem for the Field Supervisor, it does represent an opportunity for something to go wrong. In a similar fashion, attribute rotations on a paper questionnaire seem easy to administer. You start at the attribute that is checked and follow through in order. But be sure to put a strong reminder at the bottom of the attribute list because someone somewhere will forget to ask about the attributes at the top of the list.

The computerized interview can eliminate the need for multiple versions of a questionnaire when rotating concepts, products, or commercials. The computer randomizes the pattern for every interview. The task of the interviewer is simplified because he or she is presented with stimuli listed on the computer screen. The issue of randomizing attributes is handled simply and efficiently with the use of the computer.

The field execution capabilities of the computer will relieve the pressures on the interviewer. This is evidenced by the unsolicited comments voiced by our field services on various projects. For example, during the past year, we conducted a taste test with children 3 to 12 years old. This project involved two cells. In the first cell, each child was to evaluate one test product and a control product. In the second cell, each child evaluated another test product and the same control product. Each child was randomly assigned to one of two cells. In addition, the order of presenting the test and control products to each child was randomized.

These products had to be heated for a certain number of minutes, then cooled before serving. We wanted to assure that the kitchen staff was ready with the appropriate product. The computer interview was designed so that after recording each child's age, the quota group assignment and the order of product presentation were indicated on the computer screen. This information was copied by the interviewer onto a sheet of paper and used as a guide by the kitchen staff.

This study was conducted in several cities across the country. We used one field service in Texas that had never done a computerized study. The staff of this field service called us and said they loved the study and the kids really enjoyed doing the project. The field service explained that they liked the interview because everything went so smoothly. The kids were a pleasure because they remained interested in the interview.

Another project offered a totally different set of potential problems that the computer easily smoothed out. We wanted to find what two flavors children wanted together in a product. For example, suppose the product was a two-stick popsicle. The question was: What two flavors should be put together on the same popsicle stick?

The children were shown disks shaped like the product. Each disk was colored and given a flavor name. Approximately ten flavors were involved. The children were asked to decide which two flavors they would most like together in the same product. This process was repeated three more times to identify each child's second, third and fourth favorite pairs of flavors.

Up to this point, the interview was fairly straight forward. Next, we told the children to look over their four pairings. They could change their second favorite pair of flavors from orange and tango-tango to chocolate and lime. This opportunity to change any or all of four flavor pairings could have spelled disaster if the study was done by paper. A flap might have been used to record the four flavor pairings about which the rest of the interview would be conducted. When the questionnaire referred to the first chosen pair of flavors, the interviewer would have to look at the flap and read the two flavors listed. The possibility also exists that the interviewer might make a mistake and read the second pair of preferred flavors from the flap when the question called for the first.

By using the computer, there was no room for error on the part of the interviewer. The computer kept track of the flavors chosen in each pair. In fact, when the children changed their flavor choices, the computer kept track of this information. Therefore, in a follow-up question, the correct flavor pair was automatically inserted into the question by the computer.

The client on this project had been at the briefing and seen the first day's interviewing. After the study was completed, she expressed her appreciation of how smoothly the computerized interview had gone and her concerns about what could have happened if the interview had been conducted with traditional paper methods.

As you can see, there are a variety of ways in which a computer interview can simplify the task of the interviewer and make the situation more enjoyable for the respondent. I want to give another example. This is probably the most complicated central-location design that we have used to date.

This study involved testing the appeal of new dolls among children 5 to 11 years. Some of the factors that made this a particularly tricky study from a field execution standpoint were:

- There were ten items in the line of dolls being tested. The child selected the favorite and second favorite. The interview then involved questions about this favorite doll.

- There were seven controls. Four of the seven controls were used with boys and four of the seven controls were used with girls. Each child saw two of the four relevant controls. The two controls were randomly selected for each respondent.

- Each child saw a videotape for the test line of dolls and the two controls.

- The mother of each child was also interviewed. The mother was exposed to the child's favorite test doll and the same two control dolls that her child saw.

A great deal of activity and planning would have to go into each set of mother and child interviews. Needless to say, this project went off very smoothly. In fact, we have used this same design a number of times in many cities, and the comments from the clients and the field services are always favorable. This favorable response is because the computer breaks each step of the interview into its simplest parts, presenting only one action or question at a time for the interviewer to be concerned about. Irrelevant material never gets in the way of the interview.

One of the things I particularly enjoy about the Ci2 System is that many different capabilities have been built into it. An application of each capability is made apparent in the manual. However, the real creative challenge comes in applying the system's capabilities in ways that were never originally dreamed of by Sawtooth Software to meet the demands of the research design.

For example, the program has a command designated "NOA," which can be used when "no answer" is needed to a question. The manual indicates that you should always use this command in the last question of the interview. Our last question usually reads: "Thank You Very Much For Your Assistance." The "NOA" command also controls how long a question remains on the screen. We program the last question to stay on the screen five seconds. This gives the interviewer time to read the last question without rushing.

We had the opportunity to use this command in a completely different way. We did a name test with both mothers and children. We exposed five names to each respondent one at a time. One of our key measures was how well each respondent remembered each of the test names. The recall of a name could be affected by how long they were exposed to it. Therefore, we exposed each name for only 10 seconds. We accomplished this through the "NOA" command.

The computer instructed the interviewer to show a test name to the respondent, then immediately press "1." The following message appears: "The computer is timing ten seconds. The respondent should be allowed to look at the card until the next screen comes up." The next message that the interviewer saw read: "Time is up. Please take card back from the respondent. Now press "1" to continue."

Controlling the amount of time that the respondent saw a name was made possible by using the "NOA" command in an unusual but productive manner. Similar creativity can be applied to the other capabilities of the system. I venture to say that virtually anything you want to do can be accomplished with the computerized interview with greater control over the field execution of the project than can be attained with traditional paper-and-pencil approaches.

The final area relates to self-administration of computerized interviews by consumers. A common perception is that people can be in front of the computer and answer the questionnaires by themselves. Some people hold that perception in mind, while privately questioning the workability of such an approach. They might have concerns because they themselves have felt intimidated by a computer some time in their lives.

Our experience with adult female consumers suggests that a middle ground between self-administration and interviewer control is a practical course in many instances. We observed many respondents when we first started using the computer. We also spoke to several clients who had conducted studies using similar systems. The first study we conducted was set up so that the respondent sat in front of the computer and answered the questions on her own. An interviewer also sat next to the respondent to answer questions about the study as well as to ask and record the open-ended questions.

We observed through a one-way mirror that the women were initially hesitant to press the keys on the keyboard to record the answer to a question. With a little encouragement from the interviewer, they overcame their hesitancy. We also noticed that respondents spent an inordinate amount of time staring at the computer screen reading the questions and answers. This may have been due to poor eyesight or trouble interpreting the written word. For example, when a series of attribute ratings were presented, the respondents would not always notice that the rating scale remained the same, and only the attribute phrase changed. The interviewer would point this out. The result was that a 10-minute interview took 20 minutes.

Nonetheless, there are times when the self-administered approach offers a substantial benefit in allowing respondents to candidly express their opinions. This is probably the case with political polls and surveys on other sensitive subjects.

In a study of female personal hygiene products, we developed an interview approach that overcame many of the problems we had identified. It involved the use of an interviewer for parts of the questionnaire, while the respondent answered independently at other times.

At the beginning of the interview, the interviewer was seated directly in front of the computer. The respondent sat next to the interviewer and could see the computer screen. The interviewer read the first few screens of the interview to the respondent. These screens explained to the respondent exactly what was taking place and made her comfortable with the interview and the computer.

Some of the introductory screens read: "Today we are going to talk about Personal Care Products. As you can see, this computer has the questions and by hitting a key on the keyboard your answer goes into the computer. I will let you sit at the keyboard and answer some of the questions yourself. But first, you can sit next to me and watch how I read the questions and put in the answers."

The next screen determined the respondent's occupation and read: "Look at the screen with me. You can see the question and the answers. I'll read the question out loud. You read it on the screen. Then tell me the answer. I will press a "1" or a "2" which are shown next to the answer choices." Through these questions the respondent was drawn into the interview and became accustomed to looking at the computer screen and hitting the keyboard.

The next question asked about the respondent's age. The interviewer read: "In this next question you can put the answer into the computer by yourself. Read the question out loud. Then press the number on the keyboard that you see next to your answer."

Next we told respondents how to go back and change their answers. Some usage information was gathered with the interviewer reading the question out loud and recording the answers. The interview continued this way for five minutes. During that time, the respondent was shown several concept boards and several open-ended questions were asked and recorded by the interviewer. Then we moved into the attribute battery where the respondent took over. She sat in front of the computer, read the questions and recorded the answers herself. The interviewer was no longer needed until it was time to say good-bye to the respondent.

The field reported that the respondents really enjoyed participating in the study. In addition, the length of the interview was what we had anticipated.

Not every respondent requires the kind of support and encouragement that we offered in this study. Professionals may be more comfortable than female heads of households. However, I think consumers may benefit substantially from this type of warm-up.

In summary, I perceive that in computerized interviewing procedures:

- The computer simplifies the instructions and procedures that the interviewer has to follow, resulting in improved administration of the study by the field personnel.

- The capabilities built into the interviewing package (or that can be developed for it) make it possible to handle a broad range of tasks and study designs. The programmed questionnaire can be just about as creative as the people designing the research study want.

- A computerized interview can be conducted on a self administered basis providing that the respondents are comfortable with using a computer. This comfort level may be developed from a respondent's personal or business experiences. Or, the researcher might build into the interview a procedure for making the respondent comfortable with the computer interview.

COMPUTER SURVEYS BY MAIL

by
Harris Goldstein
Trade-off Research
Encino, CA

I want to walk you through the case history of a recent study, which is illustrative of computerized research conducted via mail. I hope you can benefit from my experiences. When you do a survey through the mail, your respondents must have access to the equipment. The case history I'll describe was done for Printronix, a $150 million computer printer manufacturer located in Irvine, California. An article about the study from the client's perspective appears in Quirk's Marketing Research Review (March, 1987, p. 22).

My client contact was the senior vice-president in marketing and sales; the company had no researchers. It took me a month to arrange this meeting. When he finally saw a demo, he said, "My God, I'm cancelling a flight, let's have lunch." Upon seeing what the Ci2 and ACA Systems could do, he realized that he could learn what his customers really wanted and what they were willing to pay for it.

His objective was to learn everything he could about printers in the price range of $600 to $3500. He wanted to use the study for design and sales promotion. He also wanted price elasticity information. For example, he had a $3500 printer that had been on the market for four years. They sell 500 a month. He wanted to know what would happen if he lowered the price to $2700. It was a legitimate question, and he saw that trade-off analysis would be very beneficial. Conjoint is the perfect tool in an engineering company.

Who is the qualified respondent? He is the decision maker or the decision influencer, who differs by organization, but ultimately, you're trying to find the person who specifies the brand. In larger companies with sales of a billion dollars, he may be a person who buys peripherals rather than the actual computers. We needed a large sample, because the client wanted to look at several market segments. Since budgets were tight, we decided to send the survey via mail. I'd been chomping at the bit to experiment with sending the ACA and Ci2 System program through the mail. This was a great opportunity because most business organizations in the U.S. use IBM-compatible computers and printers. The study design fit the Ci2 and ACA Systems very well.

The questionnaire started with a Ci2 System interview, followed by ACA. The first draft of the questionnaire took an hour to answer. My client wanted to learn many things, but the questionnaire needed trimming. When you design a questionnaire, include all the information, then sit down with your clients and have them make certain trade-offs. It's hard to determine early on what they'll give up to get. This is true whether you're sending it through the mail or administering it in a controlled environment. By mail, you have no control over how the person will answer the questionnaire. Will they get tired or be interrupted by a telephone call? You have to trim it as far as possible. It takes determination to decide what's critical.

We started with 18 trade-off attributes. By forcing ourselves to prioritize, we were able to give some up and cover others in the Ci2 section. If you're a printer manufacturer you care about whether an application uses single sheets of paper or six. A desktop PC uses one, but at Avis Rent-a-Car, they must print six-part contracts, and the bottom form has to be as clear as the top. You need certain software configurations, applications of print heads, print speeds, motor drives, and transactions batches.

We ended up with 12 attributes and 50 Ci2 System questions consisting largely of background information such as the size of the company, the type of industry, the kind of hardware equipment owned, the budget in 1987 for computers of various types, and the number of printers purchased in total. We learned who our respondent was, and whether he was responsible for a department or a branch. We were careful as we projected and weighted various volumetric data to get information about the whole company or a particular branch. Many people don't know what's occurring in a billion-dollar company; they may be buying for a division in Denver of a company headquartered in upstate New York. So we determined their limits of responsibility. The client wanted everything in one questionnaire. I did not give it to him. Instead, I convinced him we needed at least two versions, one that handled printers in the $600 to $1,400 price range, and another for more expensive printers.

It is critical to focus respondent attention on specific needs. These respondents could purchase a $3500 laser printer one week, and six or seven desktop, dot matrix printers the next. They needed to focus on a competitive frame of reference. We asked them to think about the next purchase they'd have to make. What is it? What hardware will connect to it? Where will it be placed? We wanted them to visualize whether it would be housed in a computer room, the secretarial area, an open space, or an executive's office.

It took longer to write this questionnaire than any I've ever written - two months. The day I took it to the post office, it had to be right. Once in the mail, I couldn't call my field agency and ask them to change a field disk.

When we finally thought the interview was ready we did six one-on-one interviews with qualified respondents. I also had four members of the client's marketing team answer the questionnaire and sign off. This insured everyone was comfortable and involved.

The mail-out sample size was 1600. We started with a sample list of 3500 sites. One critical step to insure a good response rate from mail studies is to call everybody to make sure that they are qualified. We asked them what kind of printers they would most likely buy next, so we could determine the proper questionnaire version. We also asked if they had access to IBM-compatible equipment and if they would be willing to do the survey. About 85 percent of the qualified people we contacted agreed to participate. Our sample included MIS directors, marketing executives, and CEO's from one million to more than billion dollar organizations. We had a very diversified list and wanted to represent all types of organizations.

We pretested 100 surveys during December, which is probably the worst time of year to do a pretest. Within three weeks, we had a 40 percent response rate. We mail-merged the cover letters to customize addresses. Rather than "Dear Executive" or "Dear MIS Director," each letter had the recipient's name. We sent the disk to government employees, educators, local organizations, manufacturers, the hightech industry, and hospital administrators.

We also factored in some of the client's lists and identified those people. When we weighted and analyzed the data, we would have a nationally projected sample as well as the ability to isolate special interest groups.

Before mailing a computer-interactive questionnaire, I recommend testing the survey on as many monitors as possible. I own a Leading Edge with a color monitor, an Eagle Spirit, and an IBM XT, and I tested it on all of those. I made sure it worked on a Compaq, an AT&T, and a Zenith laptop. Yet I still received a few notes saying, "We couldn't read a couple screens." Make certain it works on monochrome screens, because most screens are monochrome.

You may want to ask if the respondent uses 5 1/2" or 3 1/2" diskettes. The 3 1/2" format is becoming very popular. We use a professional disk duplicating service. You cannot be too careful.

A problem with mail surveys is that you need a questionnaire number to keep the files straight in order to tab the data. We had 1600 disks, and I had to number each one. Since there were two versions of the questionnaire, we coded one version with three-digit respondent ID numbers and the other with four-digit numbers. You have to be careful of making mistakes when you number that many disks. We also key-coded the mailing list to keep track of returns.

I want to read the cover letter, because it's important to be very honest with people. I don't know if this is the definitive letter, but it worked for me:

> Dear Mr. Stevens,
>
> You have been selected to represent your company in an important survey about computer printers. The information you provide will help a major manufacturer design better products, establish competitive prices and offer the best promotional programs to customers. Naturally, all of your information will remain strictly confidential. No sales calls will result from your participation.
>
> We are attempting to obtain information from people who actually purchase or influence the purchase of printers at their companies. If you do not qualify, please pass along these materials to the person in your company, department or branch most responsible for

buying the computer printers. This survey is different. It's a computer interview. The enclosed floppy diskette is the complete interview. Just pop the disk into the "A" drive of an IBM PC or compatible, turn the machine on and it's all automatic. Back-up booting instructions are on the disk envelope.

Please accept the attached dollar bill as a token of our appreciation. Note: this survey takes 20-25 minutes. Please do it all at one time.

We mailed out 1500 questionnaires between January 10 and January 15. Within a week, we had more than a ten percent response. Half of the returns came back within two weeks. We had a 50 percent usable response within four weeks and still get questionnaires back. When I get back to my office next week, there'll be three or four more diskettes with apology notes. People love this, especially data processing people. They compliment you on the software and send their own pet programs.

We looked at the utility scores of the first 100, the second 100, and the third 100 of each version. They were virtually identical, regardless of whether we used data from 100 or 300 people. The data held together.

One problem of doing individual questionnaires is the inconvenience of merging the data. Be careful when you check them in that the files are full and you're not mixing versions with different numbers. The system cannot remove a single respondent. You have to go back to the beginning if you make a mistake.

The cost of the mail survey was substantially lower than a standard personal survey. In the future, it would be wise to include a reference guide of pictures and various stimuli for people to look at. You can't send them a sample of a popsicle, but you can do many other things.

The combination of phone/mail/phone is a worthwhile blend. If you want the value of using mail as well as phone, you can send them some of the rating scales and attributes, then call to ask if they'll participate. Send it to them, then call them.

Computer surveys via mail work very well. The post office hated me, but I was euphoric knowing that people enjoyed the survey and returned the disks.

COMPLEX INTERVIEWS WITH LAPTOP COMPUTERS

by
Joel Gottfried
National Analysts
Philadelphia, PA

At National Analysts, I'm using laptop computers in a very large and complex study in direct response to the needs of our client. Briefly, I'll describe the study, its history, scope, the selection criteria used for the laptops, their purchase, and arrangements for use. As for software, we're using the Ci2 System with modifications made to it through acquisition of the source code.

The study itself is the 1987-88 nationwide food consumption survey. It's sponsored by the Human Nutrition Information Service of the United States Department of Agriculture. This is the seventh national survey, which the government does once each decade to gather statistics about food and nutrition in the United States. This is "the" definitive study on food and nutrition consumption in the United States, and the fourth consecutive study in which National Analysts has been involved.

Previously, these studies raised two concerns. One is the quality of the data. Being a very time consuming and lengthy study, the client was concerned that the homemakers' responses to the questions be correct. Second, it took many months to collect and analyze the data before they could be published. The client's business was to quickly publish the results and make the database and reports available for use.

The goal of the study is to measure the current level of food and nutrient consumption in United States households and also for individuals. They look at food and nutrition consumption in two ways. At a household level, it's food coming into your kitchen and leaving your kitchen in a certain time period. That food might feed not only yourself and your family, but also guests. How much food is available for people to eat? Who are these people and how much are they eating? At the individual level, it's tracking one person for three days to measure food eaten at home or away from home. As Americans, we eat away from home more often, which becomes increasingly important.

The study is large and complex. It covers a full year. The interviews are conducted every day of the week through all seasons of the year. We have no seasonal problems or day-of-the-week problems. We cover two national probability samples. The first has 6000 households to represent the United States. The second will interview 3600 households with an income restriction to look more carefully at households that are 120% or less of the national poverty level. In order to conduct these nearly 10,000 interviews, we've hired 200 interviewers. That means one laptop per interviewer, unless they work closely enough to share.

The training task is enormous and includes not only understanding the laptop and software, but also understanding enough about food consumption to conduct an interview. Laptops will make it easier for the interviewer to conduct the survey than the old-fashioned paper-and-pencil method. It takes five full days, that's 9 am to 5 pm plus homework, for the interviewers to learn enough to conduct the interview. It's a very long interview, and I am amazed that people cooperate with us. The major reason they cooperate is patriotism; I can think of nothing else.

The average length of the household interview using paper-and-pencil methodology is two-and-a-half hours, and about an hour per person for the individual food intakes. Some do not have to be done in the presence of the interviewer, which reduces the time slightly. We ask questions in the interview about the household, about the people in the household, and most importantly about the foods consumed. Each food observation has to be coded and analyzed. We ask how much, the form, the price, all the basic information about the food, then convert it to its nutritional equivalencies on 30 different nutrients that form the database.

We expect over the course of a year to have more than half a million food observations from our 10,000 households. We expect to have more than three-quarters of a million individual intake observations. So "large" is an understatement. The problem in the past has been: What do you do with these half million food observations? Food comes in so many forms and fashions. If somebody said, "I had a can of tuna fish for lunch," was it a four-and-three-quarter ounce, five-and-three-quarter ounces, or six-and-a-half ounce can? Was it labeled low sodium? Was it packed in water or oil? In the past, all they could say was, "I had some tuna fish in a can." Then we had to use the most common can size to derive the nutritional data.

We had the opportunity of using computers to customize each question. The computer screen listed the particular choices related solely to that food. That was the big breakthrough in this study. Once we decided to do the interview on laptop computers, we had to decide what hardware and software to use. The hardware decision was less time consuming and easier to resolve than the software decision.

We drew up a list of two criteria that we used in selecting the laptop computers. It had to have ample memory to run the program we would be using. It had to have the right kind of disk and number of disks. The processor had to be IBM-compatible. The screen had to be big enough and legible enough so that our interviewers could read it easily in all kinds of lighting conditions. The interview usually is conducted in somebody's kitchen, and the lighting conditions are not necessarily good. The keyboard had to be full-size so that it was easy to type in all the answers.

Our interviewers are not Olympic athletes. They are generally diminutive in nature and could not and would not lug around very heavy machines. Two years ago this study would have been impossible, because there were not machines light enough. We insisted that the computer be less than 12 pounds.

As important as the weight was how the computer was powered, especially in a poverty sample. We could not count on electricity being available, and we certainly could not count on people's willingness to let us plug into their outlets, even though the amount of power used is negligible. Battery power was required which had to last for a two-and-a-half to three hours. Some interviews would even take four hours. For a safety factor, we needed at least a five-hour battery life to conduct an interview.

Just one interview is conducted per day. The respondents are recruited and a week later they are interviewed about the previous week's food consumption. We set them up a week ahead of time. The availability, serviceability and carry-ability of the computer were all very important. We insisted the computer be shaped in such a way that it had a handle to make it easy to carry in a case with a shoulder strap.

The computer had to be reliable and durable. We had no information on that, since the machines were so new. We planned to use the Ci2 System to conduct the interview, so the computer had to be able to run the Ci2 System. We reviewed the then current marketplace and found 12 laptops that seemed plausible. We quickly eliminated seven because they didn't meet all of our minimum standards. We found five that met all the minimum standards: the Data General One, the Kaypro, the IBM, the Toshiba, and the Zenith. I had my office filled with laptops, running test programs, and everybody was looking in and making comments. I brought them home. I carried them around. Once while visiting my inlaws, my wife carried one in the car on her lap until the battery ran out. We tested them constantly. We chose the Toshiba T1100, although it wasn't the lightest, didn't have the longest battery life, and wasn't the cheapest. But it was second in all the characteristics.

In our proposal to the government, we requested 235 laptops for 200 interviewers. We wanted some back at the office to use if one broke. We were granted 210 laptops, which we acquired directly from a Toshiba distributor, rather than a retail outlet. We made special servicing arrangements with the distributor, so that we could be guaranteed a fast turnaround in service. They have set up a special phone number for us for telephone diagnostic support to determine whether the machine has to come in to be repaired. They've guaranteed us 24-hour turnaround, which means swapping machines at their end if necessary.

The technology improves yearly in hardware, and although it took a great deal of effort and energy to decide on the hardware, it wasn't very difficult to find a machine that met our standards. If we were to review it now, we'd probably pick a different machine. The technology is constantly getting better and there's nothing but bright horizons for this kind of interviewing.

For the software, we had several choices. We looked at several packages and were not happy with them. We had used the Ci2 System in less demanding circumstances very successfully and thought that formed a base on which we could expand. We instantly identified four or five things that had to be done in this interview that the Ci2 System could not do. We spoke to Sawtooth Software and acquired the source code to the Ci2 System. Then we broke our development team into two parts. The first part was to create what we called "enhanced Ci2," which added features needed for this survey that would be useful in any survey. For other features specific to this survey, the knowledge we gained in doing it would be useful to apply elsewhere, but the actual programs we wrote wouldn't be. We created 36 new Ci2 System commands, which made a substantial increase in the size of the program.

One particular point is very important. This is no longer, by any stretch of the imagination, a self-administered product. This is for a trained interviewer to administer to the respondent. All of these commands are done only in that arena.

What did we do to enhance the software? The modification that took the most time was called "edit mode," which is a general, nondestructive means to review and alter any previous answer. When you're in edit mode, the previous answer is displayed in a special box in reverse video. We programmed one of the function keys on the keyboard to "accept previous answer" to avoid typing it again. If you change an answer that changes a skip pattern that presents questions that weren't answered before, the system insures that they're answered before you escape from edit mode. Similarly, if you pick a path in which answers that had been answered previously are no longer applicable, it automatically erases those.

We also needed the capacity to handle repeating questions. We wanted to ask the same set of questions: How much did the food cost? Where did you get it? We wanted to ask these same questions repeatedly. We couldn't make a separate set for each potential food item because there could be hundreds. We implemented a way of repeating the same set of questions. This is generally applicable in household interviews where you ask the same questions of each person. We had that in part of our survey as well, and would ask the questions, then come back and loop again. The answers are saved in a separate file from the normal Ci2 System answers, and the number of repetitions is limited by the disk space. Each pass through these questions is independently filtered so that you don't have to ask everybody in the household. For example, if it's a person-based question, you could ask only women or women more than a certain age on each pass through the questions.

Unlimited comment ability is another issue. We have a policy that our paper questionnaires have wide margins so that our interviewers can scribble in notes. On the computer they can hit a special function key and a screen pops up for them to type in any comments, which are stored in a separate file. We also issued a command called DEC for flexible entry of decimal numbers.

On-screen grids were also implemented. Many questions are best presented with a grid and we left the grids intact. Previous answers are maintained on the screen. You can skip horizontally and vertically, depending on which way the questionnaire normally would flow. A few other modifications are special "don't know" and "refused" keys. These, again, are function keys we programmed. If the respondent says, "Don't know" or refuses to answer a certain question, the interviewer hits the appropriate key. The advantage is not having to include them in the answer list. It permits, yet does not encourage, "don't know" and "refused" answers, and inserts a standard code for each that simplifies programming.

These were features the client and survey absolutely had to have. The last two were ones which we could not live without as programmers. The Ci2 System lets you set a range of answers that are permissible. We found circumstances, especially in our repeating questions, where we did not know ahead of time what the permissible range was. We set up a new construct in the Ci2 System so that the programmer could vary the ranges. We also improved flow control to back up to a different target each time the question was asked, depending upon the previous answers. This allowed us to implement individual modules that we could call up without the fear of not being able to return to our starting place.

That was our enhanced Ci2 System to which we added special functions for database retrieval. There's a large USDA data base. The compressed version that we put on the diskette is about 200,000 bytes. It has 3,000 food names, 5,500 forms, over 11,000 units. We have to pull information from the database to the screen to display the food and form being discussed. It comes up as quickly as you can look at the screen. Previously, interviewers would answer "other" and type an open-ended answer. The coding task back at our shop was monumental. If we give them the complete list of foods that are currently available and known to USDA, then the number of "others" specified will be reduced significantly. The coding task should drop dramatically and we will be able to publish the data in a much more timely fashion.

COMPUTERS AND HARD-TO-INTERVIEW RESPONDENTS

by
William Tooley
IMR Systems, Inc.
Des Moines, IA

Everyone has had experience with hard-to-interview respondents. Computer interviewing has helped us to cope with this type of respondent in special ways. IMR Systems specializes in verbatim analysis, which gives us insights into how respondents behave, think, and feel. I feel privileged to share my experiences with you.

One of the first steps in any research program is the use of exploratories to develop a list of attributes, perceptions, or benefits that are important to the market. Although group discussions are often used to prepare these lists, individual in-depth interviews yield a greater variety and a longer list of important attributes. The individual interview gives a better sense of how attributes pertain to different types of respondents. Our systematic verbatim analysis provides data that allow advanced statistical techniques. We will use factor analysis of verbatim data, for example, to look at groupings of attributes that cluster naturally. This will help select the most representative ones for inclusion in an ACA or APM System interview.

There is interest in better ways to handle verbatims. At IMR we have developed a computer-aided procedure that works well, perhaps better than other systems currently on the market. We call it the Natural Language Text Analysis System (NLTAS). This system organizes the coding task in a systematic manner; helps the analyst make coding decisions by showing all the usages of the same word on a single screen; executes coding judgments across all remaining interviews; recalls previous coding decisions for comparison; shows frequency counts of expressions coded within each group; provides supervisory modules where categorization can be reviewed and changed with a few keystrokes; prints an audit trail of every coder judgment; outputs a data file compatible with most major tabular and statistical packages; and saves an automated codebook that can be read by a computer in the next study.

Since we have a system for learning about people in depth, we have been encouraged to learn more about hard-to-interview respondents. What makes a respondent "hard-to-interview?" We all come up against respondents and situations that make interviewing difficult. For example, low incidence rates, uncooperative respondents, difficult times of the year, and low involvement categories are facts of life in our business.

Let me describe several types of people who may be considered hard-to-interview respondents:

-Chief Financial Officer of a Fortune 500 company

-a respondent who is called 15 to 30 times a month

-visitors at a busy trade show who have a limited amount of time

-respondents whose interviews are regularly converted to a sales pitch

-business buyers of complex programs or services in a category that changes rapidly and is a hotbed for research

-consumers evaluating an item not usually found in their choice set

These may not always be hard-to-interview respondents. I'll describe a few of these examples further and describe how computer interviewing helped us handle them.

Consider average consumers shopping at the grocery store. They may be brand conscious, coupon oriented, particular about taste, generic price shoppers, or want certain size or microwaveable packages. Where does nutrition fit into these choices? With the ACA System, we were able to include sources of nutrition at a number of attribute levels in the grocery decision model. The consumer had little difficulty trading off various sources and levels of nutrition as additional elements in an already complex choice set.

Crop farmers are called 15 to 30 times a month during the winter. They expect their phone to ring two to three times a night. Many are still willing to participate. They seem to believe it's their duty to do so. Some will answer any questionnaire. But is it thoughtfully answered? Do they think about their answers? They know they have to tell you something, and they're very interview savvy. Last summer we traveled across the Midwest visiting expos from VanWert, Ohio to county fairs in Iowa and Illinois to the Sioux Empire Fair in South Dakota. We contacted crop farmers and asked them to spend 30 to 45 minutes at the computer answering questions about their decision model for certain products. We paid $10 to attract them, and they sat in the hot sun, under a tent, or in the field at the computer.

These busy farmers may ordinarily be tempted to give answers with little thought. The ability to get a thoughtful interview and the willingness of the famers to become involved with the ACA System made this study successful. They thought about every decision they made on every screen. Often the wife was present. The farmer would ask her, "What do you think? You're the one who buys this stuff." He'd say, "Good," then push the button and proceed.

The client also asked us to do a test. He had done conjoint analysis in the past with limited numbers of attributes and with little attention to the paper-and-pencil task. The model would gyrate fairly wildly. With thoughtfully done, full-profile conjoint data, we were able to predict a 63% market share using a simulation model. From survey data, we knew the marketplace had a 66% market share last year.

Another group of hard-to-interview people are busy company decision makers who receive numerous calls from salesmen probing their buying decisions. For most decision makers, the decision criteria we present can often seem nebulous and inconclusive. We told these decision makers that we were running a computer model (the ACA System). By going through the model, they would be given an assessment of their important buying criteria. They were attracted by the opportunity to be involved with an informative computer model that would help them face their own complex decisions. With no incentives offered, most took the time to participate and spent 20 to 30 minutes on the phone.

When done thoughtfully, the ACA System offers busy professionals a new perspective on information they provide. Assistance of this type (helping respondents to review a complex decision process) can be greatly appreciated in the information-hungry world in which we live.

In conclusion, computer interviewing is not just another way to administer an interview. Instead, Sawtooth Software has provided a whole new tool to communicate with market perceptions and fundamental decision criteria.

Research can now achieve finely-tuned results using quantitative methods, equal to those achieved using in-depth exploration and probing. This is true at the individual respondent level as well as the group level.

Computer interviewing offers a unique and powerful approach to systematic respondent data handling and analysis. This approach provides hard-to-interview respondents with the decision support and feedback they need, and as a result, offers the researcher access to better decision modeling data.

We are reminded of Marshal McLuhan's statement: "The medium is the message." Computer interviewing creates an entirely new medium for each communication. The message is that marketing research as we know it will be very different in the future.

GENERAL FOODS FACES THE COMPUTERIZATION OF PERSONAL INTERVIEWING

by
Audrey Bowen
General Foods
White Plains, NY

I am very proud to represent General Foods in this leading-edge seminar. As we have already heard, we at this seminar surely are the leading edge. Not only are we the trendsetters, but also we are in the unique position of setting up the industry standards for computerized interviewing.

General Foods has been in the computer interviewing business since the turn of the decade, when we started using computerized telephone interviewing. Being well aware of the efficiencies involved in computerized telephone interviewing, we had been waiting for this capability to open up for personal interviewing.

As we have been hearing at this conference, this has become a reality, and with the same efficiency, cost effectiveness and speed, not to mention the accuracy that had become the way of life with computerized telephone interviewing.

We have been doing computerized personal interviewing for about a year, but did not convert without a lot of pain and anguish. We had questionnaires that were never recorded, a day's work erased while downloading data, trouble with IBM PCjrs, telephone lines that went down, etc. But we were still willing to gamble because we believed sincerely in the pay-off.

We gradually set up standards to help us keep our sanity as well as to keep up with current developments in the field:

Lead City - We feel strongly that a lead city is necessary to iron out the bugs before the questionnaire is released to all sites. This lead city interview ideally occurs two days before the actual field, or at a minimum, one day with one of us in attendance to make on-the-spot changes and decisions.

The lead city can be staffed with only one or two interviewers depending on the incidence. The main purpose is to insure that the questionnaire is 100% workable. The completed questionnaires in this lead city may or may not count toward the quota, depending on the changes made.

Self-Administration - Business, especially the data collection phase of our industry, is facing serious labor shortages. In fact, this shortage is hitting all the service industries. How many times have you passed a fast food restaurant and noticed the help wanted sign? According to a recent issue of Fortune, we will have more active workers than dependents by the turn of the century. The middle-aged, baby boom folks will create a labor shortage. This, added to the fact that millions of women have left the home and part-time jobs to join the full-time market, will increase the severity of the problem. Statistics show that today 53.7% of women are in the work force, and of these, 62.8% are mothers with children under 18 who used to make up a large part of our labor force.

Recognizing the impact that this labor shortage is having on interviewing, General Foods has chosen to go the self-administered route. We program our interviews to be self-administered by the respondent. The field service still provides the screening interviewers, but inside the site, it provides only one attendant to handle two or three computers. A key to being successful in the self-administered mode is to control the number of open-ended questions. Obviously, the fewer open-ends, the easier for the respondent, and the less interaction with an interviewer. We still believe that the response to open-ended questions should be recorded by hand by the interviewer; otherwise much of the flavor is lost. General Foods is addressing this issue by not asking open-ends on a number of studies, especially on established brands.

A great benefit to this self-administered mode is the fact that _each_ respondent in _each_ city and in _every_ site sees exactly the same question, as we want them to see it. There is no room for interviewers to paraphrase or ad-lib.

An added benefit is that respondents become involved in the process. They enjoy it and even feel that the interview is about 40% shorter than it is.

Review of First Day's Data - In a paper questionnaire, missed rotations, skip pattern errors, and an illogical question are easier to detect. In fact, in some situations, they just jump out at you. On the computer these errors are more subtle. One must go through each skip and series of questions, but even then a loop or two may be missed. To guard against this, the first day's data should be downloaded and scrutinized for logic and consistency. We can discover rotation errors as well as skip errors in time to correct them for the major part of the study.

Cost Efficiencies - The obvious cost savings come in the form of both personnel and paper. On the personnel side, we use fewer interviewers and have no need for editors, since we have no paper questionnaires to edit. On the paper end, we eliminate the juggling of questionnaires, the collating of the attachments, the validation forms (respondent's validation data are input during the interview), not to mention unpacking and counting, and packing and shipping from each field site.

Other savings are not as obvious: All the printing and the production work in rotation, product coding, and interleafing; packing lists and checks against these packing lists before the client ships; shipping out and then shipping back (one, two, or three times); check in; and finally, keypunching with those inconsistencies that tab houses always discover.

All of these costs are not necessarily savings from a field agency point of view, but a considerable number are. At General Foods, we do not expect our field bills to be any higher than they would be if the interviews were conducted on paper. In fact, we expect them to be lower just because of the built-in efficiencies of computerized interviewing and self administration.

In summary, we at this conference are the principals in a major part of the technological revolution, or should we say, evolution, taking place in the Marketing Research Industry. We will continue to become even faster and more efficient and more effective.

We represent just a small percentage of the industry, but we are definitely the leaders. But to remain leaders, we have to anticipate future needs, developments, and technological advances, then make the next steps happen.

The next steps that I foresee are:

- Touch screens to make self-administered questionnaires even easier.

- Interactive communication, for example, job specs, bids, instructions, changes, and even the final billing.

- Digitized voice receivers.

- Easy-to-access daily counts, instant toplines and clear tables for those doubting, rejecting, "yeah, but" project managers.

THE FUTURE OF COMPUTERIZED CENTRAL LOCATION INTERVIEWING

by
David Griscavage
Pepsico, Inc.
Valhalla, NY

Several factors can affect the future of computerized central location interviewing:

PC Networking

The availability of a PC local area network (LAN) and representative software will make quota control on PCs a reality. We look forward to the development of this factor since it will computerize one of the most problematic areas of interviewing research - the maintenance of quotas.

The Use of Modems

Ultimately, we would like all of our questionnaires and data to be transferred electronically via modems. We have used electronic transfer in the past and found it troublesome. The full utilization of the Ci2 System will not be evident until most field services have and can use modems, and until large scale transmission of interviewing information via modem is available.

Self-administered Interviews

An area that needs investigation is self-administered interviewing. Research should examine what differences exist when an interview is conducted by interviewer vs. self-administered. Self-administered interviewing, with its attendant lower costs, is an attractive option with the Ci2 System. Once its advantages and limitations are known, it then may be more fully utilized.

Widespread Computer Penetration and Training Among Field Services

The field service industry holds the key to computer interviewing. Their acceptance of the technology and their commitment to train personnel and to gain experience will ultimately affect the modernization of the industry. Many field services have already accepted this challenge. We look to their leadership in the future.

Computers as Bottlenecks

Computer interviewing can create bottlenecks. The ratio of PCs available to respondents should be 1:1. However, in shopping malls where traffic can vary, the number of potential respondents to computers can become problematic. The best solution is to work with your field service and have them recommend a number to use based on their experience.

Where Are the Data Diskettes?

It's very important to keep track of data diskettes from their point of collection to their first shipment to the client. We have lost no diskettes yet, and hope that this record continues.

To CUM or Not to CUM Diskettes in the Field

Our policy has been not to CUM field diskettes in the field. This process is best performed by individuals who are familiar with the process - us or one of our suppliers. As field services become more experienced with procedures, we may ask field services to CUM diskettes as a method of making a copy of the data.

WHO SHOULD DO WHAT? FIELD VS. SUPPLIER VS. CLIENT

by
David Santee
Hallmark Cards, Inc.
Kansas City, MO

The advent of computer interviewing has provided the opportunity for field agencies to offer several new services that can give them a more important role in the survey research process. The following are four such services:

1. Quality Interviewing

2. Questionnaire Programming

3. On-Site Programming Capabilities

4. Transmission of Data

Quality Interviewing

It is hoped that most field agencies are currently providing this. However, the skill level of interviewers becomes more important when conducting computer interviews.

We ask interviewers to have more skills than in the past. They must know how to load a program and terminate or restart an interview. Typing skills are also required if an interviewer is to enter open-ends.

Supervisors need skills beyond those of the interviewers. Supervisors must know how to cumulate and back up data. They should know enough about the system to troubleshoot when problems arise. Ideally, they should be skilled enough to offer suggestions to their clients on ways to improve questionnaires or the interviewing process.

My perception is that interviewers have not been adequately trained. Hallmark has talked with field agencies that advertise expertise in computer interviewing, yet have little or no experience. Many could not handle slight variations in the system or were completely stumped by minor problems.

Much of this is expected, given the newness of the system. Yet, by this time, we expect qualified interviewers who can cope with minor problems.

This conference has concentrated on analytical techniques. But we must remember the basics. If we get poor information from the field, no conjoint simulation method will give good information.

One thing that keeps me from utilizing computer interviewing more than I do is the time it takes to program the computer questionnaire. I can write paper-and-pencil questionnaires more quickly. In fact, I typically draft a questionnaire on paper prior to programming it on the computer.

Field services should offer to take a paper draft and adapt it for the computer. The field service that does this is probably the one who will field the study. An additional benefit is that the field service can handle programming problems if they arise in the field.

On-Site Programming Capabilities

For some studies, no amount of pretesting will catch all problems that may occur in the field. Or we may have a rush study with no time to pretest. Typically, the program disk is sent back to the client for him to correct and return to the field. Two days later the interviewing can begin again. This results in 1) higher costs for the field service due to idle interviewer time, and 2) a delay, which can be critical for rush projects.

If changes can be made immediately in the field, these costs and time delays can be avoided. Interviewing can be resumed in minutes.

I am not suggesting that clients give field services a free hand with altering questionnaires. Changes should be made only after consulting with the client. A questionnaire with major changes probably should be pulled from the field. But for typos, unclear instructions, missing randomization or skip instructions, on-site field changes seem appropriate.

Transmission of Data

Since we collect data by computer, we have the main component necessary to transmit data to the client through a modem. Many of us already have modems, which are fairly inexpensive.

Modems offer several advantages. They are faster and less expensive than overnight delivery services. The primary benefit is that they can transmit data frequently during the fielding of the study.

Transmitting data during interviewing allows the analyst to look at it, set up tables, and test programs. When the entire sample has been gathered, the analyst will need less time, decreasing the turnaround time of the study.

By transmitting data frequently, the analyst can monitor the variance of key questions. A sequential sampling system can be established whereby the analyst readjusts the sample size depending on the observed variance. Such sequential sampling may permit a smaller sample size, saving time and money. Or it might suggest a larger sample size. If so, we have time to find extra samples and make statistical inferences from them. Frequent transmission of data through a modem makes sequential sampling possible.

General Observations

These and other services provide an opportunity for field agencies to differentiate themselves from their competitors. If field services ever were a commodity, they are no longer. A dichotomy will soon occur: full-service field agencies that provide all services through tabs, and limited, low-cost agencies that just collect data. Each agency needs to decide which strategy to pursue.

A benefit of this entire process is the opportunity for the field agency to become more of a partner in the survey research process. Just as research suppliers do a better job when they are partners rather than order takers, field services can provide better information when they, too, become partners.

WHO SHOULD DO WHAT?
FIELD VS. SUPPLIER VS. CLIENT

by
Peter Honig
Peter Honig Associates
White Plains, NY

Much of my thunder has been stolen by Audrey Bowen and others who have spoken before me. I agree with most of what has been said. I hope my suggestions will help place us in the forefront of accelerating the growth of computer-assisted interviewing.

First, what do we expect of the field? We expect the field to follow our instructions. I hear some laughter from the audience. If you're a field person, it's probably nervous laughter; if you're a supplier, it's probably knowing laughter. For example, we did a study recently and the instructions were very explicit: When the respondent returned to the central location with any test product that had not been fully used, the interviewer was to destroy it. Last Thursday, we received a huge Federal Express shipment containing boxes of unfinished test product.

As far as Ci2 System interviewing is concerned, the field should understand that our clients may want verbatim answers. We want to print them out without editing them. Therefore, we want literate interviewers who can type responses with good grammar and spelling. We expect the field to train their interviewers properly, both in terms of how to operate the computer and the nuances of the questionnaire. We write instructions for every study that we do and like to have those instructions followed. We like to give pronunciation guides, etc.

The field should realize that we have timing requirements and firm commitments to our clients. Overbooking is something that we all have to live with. Much of it is caused by last minute changes and delays. The client should be willing to pay for delays he causes. However, with the availability of hardware rental facilities in most markets, we should no longer hear: "We don't have enough computers." Renting additional computers is a real option.

We still have to provide a reasonable value to our clients in order to sell computer interviewing. Deciding what the upcharge should be is a very relevant issue.

I have a strong commitment to computer-assisted interviewing, and I want to tell my client it does not cost any more than traditional paper-and-pencil interviewing. The data are cleaner and a better study will result. If you tell them it costs more, they'll want to use paper-and-pencil. That's a dilemma.

We realize that it is difficult to staff interviewing personnel. Good people are harder and harder to find. The analogy was drawn to the need for help in fast-food restaurants. We're competing for some of those people. Computer-assisted interviewing and the use of self-administered questionnaires should help remedy this problem.

If there's a problem, we expect the field to report it to us immediately. Don't wait until we call you to say, "Oh, yes I meant to call about that." You can prevent a great deal of pain in other cities if you let us know immediately.

Ideally, the field should have access to the Ci2 System to make minor changes in the questionnaire to prevent downtime. The perfect questionnaire does not exist. Important errors should be fixed on the spot without delay. The supervisor need not be an expert in the Ci2 System. We can "walk" them through the program by phone and tell them exactly what frame or logic instruction to change. They can do it on their screen and read it back. They can also FAX a copy or send it by modem to be checked. In the long run, field services that have knowledge of how to use the Ci2 System will get a greater share of the work.

The upcharge for the computers should be minimized. As a supplier, we do not want to charge more for computer interviewing, and if priced properly, we will all profit. Reasonable pricing will result in greater volume. We will not make as much per interview or per study, but we will do more interviews and more studies.

There are many additional benefits of computer interviewing. At the field level, computer interviewing will result in the ability to get interviewers up and running faster. Furthermore, the use of self-administered questionnaires will create greater efficiency. One interviewer can perhaps oversee three or four computers. If we promote computer-assisted interviewing, we will all benefit.

I want the field to understand IBM DOS, especially the file copying capabilities. I want them to understand the Ci2 System functions and to keep a log of any computers or disks that give problems. We had some problems on a study, and we weren't sure whether it was the program or the computer. We pinned it down to a hardware problem by keeping track of the computer on which each malfunction occurred.

The client should understand the time required to program and test the questionnaire and to make field disks. Do not expect to approve a questionnaire on Thursday and have the field start interviewing on the weekend. Sufficient lead time breeds quality. I would also like to see clients limit open-ended questions. Except for communication-type questions, open-ends tend to be over used and drive up the cost. If you know the category or product, closed-ends can be structured.

Another wish is that clients learn to accept the Ci2 questionnaire printout as a draft. This would save time and the duplication of effort in first developing a conventional paper draft for approval.

Finally, what can we the supplier do to improve the efficiency of computer interviewing? Let me list some self-explanatory items:

- Provide proper instructions for the field:
  - an overview
  - an understanding of nuances of the category, questionnaire, and pronunciation guides
  - provide instructions with specific examples

- Proper pretesting:
  - x-backs
  - setting of flags for reference

- The use of simple easy-to-read screens:
  - highlighting
  - flashing
  - color or reverse video
  - simple wording

- Proper lead time and realistic quotas
- Provide good incidence estimates for proper staff planning
- Provide sufficient field disks
- Back up all data

Some of the standards that we have set for computer-assisted interviewing studies are:

- Pretest the questionnaire thoroughly
- Provide written instructions
- All interviewers must do practice interviews
- We do all the programming ourselves
- Provide sufficient field disks
- Set a maximum number of interviews per interviewer to avoid fatigue and lowered quality
- We always use the automatic respondent numbering function
- We do not allow the use of PCjrs
- We validate

I encourage everyone involved with computer-assisted interviewing to help promote its use within reasonable pricing parameters. In the long run, we will all benefit. Furthermore, if we begin to standardize our methods of using this tool by making the Ci2 System our standard approach, the task should be even less difficult.

WHO SHOULD DO WHAT?
FIELD VS. SUPPLIER VS. CLIENT

by
Elizabeth Bradley
National Analysts
Philadelphia, PA

I work in a custom research environment at National Analysts where we do many different types of studies. The once-a-decade study described by Joel Gottfried is supplemented by the once-a-week study as well. It's the latter that I'm involved with.

When I think about the roles of the supplier, field, and client, I become confused. The roles depend on who is responsible for the questionnaire development, who is responsible for the data collection, and who is responsible for data processing and analysis. The roles are defined in association with the tasks, rather than with the particular people involved.

The primary purpose is to design a questionnaire that is correct, that works, and that collects the needed data without skipping anything. We review our questionnaires largely in paper-and-pencil format. There are many word changes and the Ci2 System is not an effective tool as a word processor. We make certain that all the changes made in the pencil-and-paper version are translated and that the additional features of the Ci2 System are incorporated when computerizing the questionnaire.

When the questionnaire is reviewed, whether by the client, our in-house people, or the field, we need to take into account whether the computer will help or hurt. The paper-and-pencil review is very important. Some people do not like the computer. Some people review questionnaires while commuting on trains. Although laptops are available, they are not the best way to get people to evaluate the words. You need to take those kinds of issues into consideration when defining roles.

During the on-line review, have as many people as possible look at it. Role playing is a very effective way to get in-house people to go through a questionnaire. For instance, assign someone to be a telecommunications manager and tell him, "You have five locations; they're all in one state; you like to use as many alternatives to AT&T as possible. Now, go through the questionnaire. Look at the skip patterns. Are you answering the same questions twice? Are you answering a question that does not make sense based on the answer that you have given before?"

That role-playing mindset is a very effective way for someone familiar with the Ci2 System and questionnaire construction to find problems. "X-ing back" is very important. Another potential problem in multiple module questionnaires in the field is respondents who need to be restarted. During the interview they may have to go to a meeting, but they're willing to finish later. Learning how to restart, especially in the middle of an ACA System module, is very important. Be prepared to answer immediately when the interviewer calls in and asks, "What's happening?"

The thorough review of pretest results is always essential, and everyone has war stories. Mine concerns using the wrong version of the ACA System. We used a demo version instead of a regular version, so instead of getting 20 respondents on a disk, we had 20 Ci2 System interviews and one ACA System interview. We were able to work with the field and turn it around overnight. We lost the first piece of the interview, but because we have an association with multiple field sites, we could get to all sites immediately overnight. Pretest checking not only of the questionnaire, but also of the data in terms of how you want to analyze them, is invaluable.

The data processing department is very helpful in this portion of the review. They are programmers who are used to looking for logic errors and writing cleaning specs. Having many cleaning specs up front in the questionnaire is a very helpful way to check for outliers. For instance, did you really mean $2,000 for that tube of toothpaste? Such errors occur in questionnaires. That's one of the benefits of computer interviewing of which you can take advantage. You move much of that process up front; normally it is at the back. People often talk about shortened time frames. You save a great deal of cleaning time if the questionnaire has been checked. We have found it just slightly shortens the total time. Involving data processing people in that up front process helps to insure quality.

In working with the field, frequent reports are very useful. In data collection with the Ci2 System, having one interviewer overseeing three PCs helps to get data faster. It also builds problems more quickly. Once-a-week or twice-a-week field reports are not enough. We need them everyday. As mentioned previously, incidence can affect the quality of the data.

Some enterprising interviewers speed things up. They go through the respondent number screen and pull the first question up on the screen, so that the interview will be ready as soon as the respondent sits down. But that starts the timer, too, so you have 300 minutes on a 20-minute questionnaire.

Field sites often have monochrome monitors. If you have designed a color questionnaire, it can be hard to read. Double check the type of monitor each time. Premature shutdowns are another potential problem. Turning off the computer at the five or nine o'clock quitting time can lose the data of the last respondent. Emphasize this repeatedly in training, even with the best interviewers, each time you take a questionnaire to the field.

"Computer coverage" means "how many computers do you need?" We have done personal interviews with the Ci2 and ACA Systems in a technical environment with communications managers and found that 50 percent filled out the questionnaire on their own PCs. More than that owned PCs, but they liked using laptops because they could use them in an environment where they were less frequently interrupted.

As for technical support, the field supervisor is not always the person to answer all questions. There are many types of PCs, and you need someone who understands them. The respondent who understands computers can run a directory of your disk and start the EXE file instead of the BAT file. You need to know how to correct such things. Someone who can answer those kinds of questions from the field needs to be available.

Conversion to different data formats may be necessary. Taking the data from the Ci2 System and putting it in your own tab package takes time, especially for the first studies. Allow yourself enough time to do it. We also use the ACA System output, which takes more creativity to put it into our probit simulator, or use the estimates from the simulator and put them back into the Ci2 data to run tabs. Whatever responsibilities you assign to whom, it's important to work together, check each other, work toward the end result, allow yourself enough time, and test thoroughly.

WHO SHOULD DO WHAT?
FIELD VS. SUPPLIER VS. CLIENT

by
Richard Miller
Consumer Pulse, Inc.
Birmingham, MI

Computer interviewing has arrived. It is growing rapidly and is definitely workable. Certain companies in the field have taken the lead in this area and are ready to conduct these interviews as their clients become ready to do so.

Our clients, both end users and suppliers, ask: Are you doing computer interviewing? Can you help me? This is the role data collection firms must play today - assisting clients with computer interviewing and providing the expertise. Here's how we can help.

About three years ago, we made a commitment to enter the computer age. After extensive research, we elected to use IBM hardware and software from Sawtooth Software. We now own 11 IBM System 36's for internal use and more than 50 IBM PC's for data collection. We own all the Sawtooth Software systems. The Ci2 System is easy to use, which is an advantage in the field because field personnel do not have great computer expertise. It also is affordable and comes with a one-time manageable cost, rather than an on-going expense. Plus, it is fantastically supported with easy-to-learn manuals and available advice and support, and it does everything you will ever want to do to alleviate errors, make a good presentation, and handle data efficiently.

In 1986, Consumer Pulse conducted over 160 computer interviewing studies. Using stand-alone Ci2 interviewing, the studies split equally between in-person interviews and telephone interviews. Using the ACA System, nearly all the interviews were mall or prerecruited. We did one ACA System telephone study.

We dealt with clients with various levels of expertise. Some were computer-sophisticated and provided us with field disks for data collection, whereas others relied heavily on our ability to put up questionnaires, create field disks, and oversee the entire data collection process.

When we are contacted for a study that involves computer interviewing our first task is to determine the level of expertise of our client and to target our sales approach and capabilities to those levels.

If our client will be sending us field disks, we must determine that the field disks will work on our IBM equipment. If the client is preparing field disks, we stress that they use only IBM DOS 2.1 for formatting field disks and transferring the system. Supplying field disks formatted with Tandy DOS or some other version of DOS causes lockups of IBM PCs or booting problems when the machine is turned on.

Many data collection firms have PCjrs. If a field site has a PC, IBM DOS 2.1, and the appropriate memory expansion program copied onto the field disk, you can conduct interviewing. If the version of DOS is other than 2.1, the memory expansion will not work and valuable days of data collection will be lost.

We next discuss how many disks to send to field sites. Most clients feel two disks are more than adequate per 100 interviews. Sending two disks, however, can create many problems. If one disk malfunctions, you are left with just one disk and one machine on which to interview.

Many clients use expensive brand-name disks, which can be less reliable than certain off-brand unlabeled disks. No matter how expensive the disk, we have received bad disks with bad sectors. This means a field site must stop interviewing. We strongly recommend that for a 100-interview study, for example, at least three to four field disks be sent. It costs less than 50 cents per disk to provide an adequate number of usable field disks.

Then there is the issue of practice disks. Although the Ci2 System has the "0" respondent number option for conducting practice interviews which is very useful for checking the questionnaire, we strongly recommend providing a practice diskette for interviewers to conduct their practice interviews. A practice disk allows a supervisor to check data more easily, and keeps practice interviews separate from completed interviews or other data on the diskettes.

If the client wants us to send them diskettes half way through the study so they can get interim reports or run a marginal analysis, we send the field disks to the client or the location that prepared them. That way, one person is accumulating and checking in the data and conducting the marginals.

If a client is interested in marginals for just one question, we use the MARG program of the Ci2 System and phone the combined data to the client. If the client is interested in more than just topline preference, we recommend he sends us a field disk called the CUM disk. This allows us to cumulate the data from the field disk on to the CUM disk, then run marginals. This information can be telephoned or the CUM disk sent directly using an overnight carrier.

Clients have asked us to make copies of the field disks after the data have been gathered. We shudder! We never make copies of field disks. Some field sites have completely lost data as a result of copying.

Many field sites own single-disk drive PCs. Making a copy on a PC with a single-disk drive requires taking the field disk in and out many times. This may result in copying data over other data. More importantly, if you were doing a paper questionnaire, would you ask your field sites to make a copy of all the questionnaires in order to send back the originals? If a client insists on making copies, the field site should have a two-disk drive machine, use write protect tabs you provide, and use the COPY *.* command. Under no circumstances should a field site use the DISKCOPY command. The client must provide detailed supervisor instructions about how to copy.

Next, we learn whether the questionnaire and interview will be self-administered or interviewer-administered. Although that decision may rest with the ultimate client, we often assist with a recommendation. Most interviews can easily be self-administered, as long as they do not contain long open-ended questions that require detailed probing.

We also discuss with the client whether the client or our data processing staff will design and program the questionnaire. If the client designs the questionnaire, we supply him with our Ci2 Formatter. This is a pad of preprinted forms developed for frames and logic that lets the questionnaire designer quickly write in questions or frames. This allows a clerk to type the frames quickly and easily.

The other side of the page provides Ci2 System logical expressions and allows the designer to record the numbers for logic instructions. The Ci2 Formatter is a quick and easy way of setting the preliminary instructions into the computer so that the debugging process can be completed. These forms are sent free of charge to clients.

Then we determine the colors to use on the screens for the interview. Color is a great help in completing questionnaires. The responses read to the respondent should be green on black, and our interviewers know to read this information to the respondents. Interviewer instructions or items that will not be read to the respondent should be blue on black. Words that need highlighting, such as evaluations, should be in acceptable high-contrast colors. We do not advise using white on gray or other low-contrast colors. Do not assume that the color shown on your computer is the same color that will appear on the interviewing machine, particularly among the new EGA color monitors.

Sometimes clients want to enter screeners as part of the interview. If the screeners contain data that need to be entered in the computer, the data should be entered by the interviewer either before or after the interview. Clients often do not allow for screeners without a completed interview to be entered into the computer. A simple logic instruction at the beginning of the questionnaire to determine the type of screener completed can then be used to create proper skip patterns.

We discuss with the client how to keep track of respondent numbers on the disks. We tell them to prenumber the questionnaires on the field disks and set the number of questionnaires on each field disk to 100. We have created a form called the "Disk Disposition Sheet," which lists the disk number and the questionnaire numbers on each disk. This allows us to indicate the name of the respondent, his telephone number, quota groups, and any notes about the disposition. These disposition sheets are returned with each disk and allow the tabulater to delete respondents who are not part of the study sample.

Then comes the issue of validation sheets. Currently, we complete validation sheets upon request. However, data for validation sheets are contained in the disposition sheets and can be included in the Ci2 System interview. Clients who require validation sheets by interviewer should consider combining data files with the open-ended file and printing their own validation sheets sorted by interviewer.

Clients often ask: How do I make a change if I have to? We have no problem making changes in the field, but we must be careful how these changes are made. The structure of the data on the field disk should not be changed. The data file may be expanded by inserting new questions, but be careful to use skip patterns that do not change the actual order of the questions. If changes are to be made, we recommend sending new field disks. This saves potential problems that copying the logic or frames onto each individual diskettes may create. However, if only the logic or the frames need to be copied, this can be done in the field. We recommend creating a "BAT" file using DOS commands to copy the file automatically onto each field disk.

Then comes transmitting the data. We have transmitted data via phone lines and modems, but it is much less expensive and more reliable to ship the field disks to the client's processing site. Although there are many individuals trained on the operation of modems, it is much simpler to accumulate the data using the CUM program of the Ci2 System, convert the data, and send the file to the client by overnight courier. Modems may take an entire day to transmit data and can result in high telephone and labor expenses with potential data errors.

The final question a client usually asks is: What materials should I send to the field? Our answer is: Send us the same information you would send if you were doing a paper questionnaire. The only change is on the paper screeners. Please include a space for the computer questionnaire number, and prenumber all the disks so we can correlate a screener to the prenumbered questionnaire.

We feel very confident about the computer interviews we have conducted for our clients. Over the years, the Ci2 System has never lost a respondent at Consumer Pulse. Our interviewing and data processing staff has easily adapted to the entire computer process for putting up questionnaires, coordinating the field sites, and interviewing. More importantly, our corporate staff provides the technical support for all Consumer Pulse offices. This is a benefit of dealing with a network that knows its equipment, rather than independent agencies.

With multiple sites, we can duplicate a problem and know whether it is a computer problem or a field disk problem. We can tightly control all aspects of a study from one central office.

Yes, computer interviewing is here. Consumer Pulse is proud to be part of this trend. It affords all of us a unique opportunity to work together in developing the research tool of the future. I hope our experience can contribute to its further development.

WHO SHOULD DO WHAT?
FIELD VS. SUPPLIER VS. CLIENT

by
Bernadette Schleis
Interactive Network
Evanston, IL

I represent the "Interactive Network." Bernadette Schleis & Associates is our parent firm. We are not an end-user, supplier, or field interviewing agency by traditional definition. Instead, we are a field direction company, specializing in data collection management. We are, then, Field Directors.

Our goal is to serve as an extension of the project teams of our clients in regard to data collection. Our clients are end-users, research suppliers, and consulting firms. While we have expertise in all data collection methods, our "Interactive Network" division was formed two years ago to specialize in field direction of computer-interviewing projects.

The more we know about our clients' needs and study objectives, the better we can function to assist them in two ways. First, to anticipate potential field problems and work to avoid them or reduce their impact. Second, to suggest intelligent solutions to field problems, rather than to dump the problems back in our clients' laps.

When confronted with the question: "Who should do what?" our company has difficulty answering because of its service orientation and philosophy. In any case, even though we choose to remain flexible and accommodate our clients' requirements, we've learned many things about "who should `best' do what."

First, allow us to select field agencies. There's not a company in the world that can meet all of its various clients' needs perfectly. Furthermore, many agencies wish to limit the scope of their work and do not want to be all things to all clients. A field agency is as good as its facilities and the skill of its local management. Our selection of agencies for the "Interactive Network" is an ongoing process. While we prefer long-term relationships with field agencies, we also understand reality. Reality and experience have proved to us that some agencies are better at some types of projects than others, and that management at the local agency level changes over time. Our support staff includes more than 45 field directors located throughout the nation. This staff helps to expand our current knowledge of the state of local agencies. Their experiences visiting agencies keeps us up-to-date.

Second, we encourage our clients to develop their own questionnaires. This gives them greater control over study design and their staff is closer to and more familiar with their needs. Also, it's more cost efficient.

Third and most important in any self-administered interview, we urge our clients to pretest computer questionnaires. This can be done internally with office staff not directly involved with the project. Test a copy of the development disk for logical skip patterns and to be sure all possible answer alternatives are accounted for. Checking a hard copy of the frames and logic is also desirable. It's very helpful for us to receive a hard copy and preliminary field disk to help in debugging. Also, after a formal or informal pretest, examine the resulting data files to be sure stored answers are there. Some of our clients even process pretest data with their chosen cross-tab and statistical analysis software.

Fourth, let us provide the additional field materials we feel are necessary for proper control. These materials include:

-Basic computer interviewing instructions outlining hardware set-up; handling and storage of disks; and how to solve equipment problems (which, by the way, are very rare).

-Monitor instructions for the person in the computer room overseeing self-administered interviews, which outline how to guard against respondents "playing" with the equipment; how to support them with answers about entering data, backspacing, changing answers; and how to make them comfortable during lengthy interviews - offering refreshments and such. Also, the room monitor may be responsible for entering quota codes or showing exhibits.

-Call record sheets tailored to the needs of the study.

-Respondent signature and log sheets for each disk to better facilitate control of quota groups and to record comments or identify problems.

Fifth, let us take care of the hardware needs for a project. Many of the field agencies we work with already own the proper equipment and others will be purchasing it as their volume grows. For non-owners, we contract with rental agencies and can negotiate reasonable fees. Because of our volume use, we receive substantial discounts. Rental fees have dropped dramatically over the last two years.

Sixth, allow us to respond properly to your bidding requests. Contrary to popular opinion, there is no one price for a computer interview, just as there is no one price for any other type of interview. Ballparks are sometimes necessary, but eventually what we need are complete specifications, since an estimate sometimes has as many as 30 cost increments. Our bids are different from field agency bids and include all data collection and field direction labor and expenses, shipping and printing, an additional 20% validation, backup disk preparation at our office, editing disks, and long distance telephone charges. Basically, we manage the data collection phase of studies from start to finish or at any point during data collection when our skills are needed. Obviously, there's no point buying what's not needed.

Finally, some other suggestions to our clients:

-If necessary, have brief questionnaire instructions to assist the room monitor in helping the respondent with difficult portions of the interview.

-Do not have field agencies make copies of field disks unless you've asked them in the past to copy completed paper questionnaires. It's unnecessary and potentially disastrous.

- Allow a refusal code in your range of possible answers to sensitive questions that does not appear on the screen and is known only to the room monitor. This is rarely needed, but when it is, it allows the interview to continue.

- Instruct field agencies in the use of copy-protect tabs when it is necessary for them to copy or edit disks.

- Use the automatic respondent number procedure and message screen.

- Provide 1-1/2 to 2 times the number of disks needed to complete quotas.

- Do not have more than 25% of any field agency's quota on one disk or 25 interviews, whichever is less, even if you have the capacity for more. The reason is the same as having multiple field shipments of paper questionnaires to guard against lost or damaged disks of an agency's entire quota.

- Do not have us run your data with cross-tab or statistical analysis software.

In closing, we use computer interviewing primarily for respondent-administered interviews. This is where both Sawtooth Software's Ci2 and ACA Systems demonstrate their greatest advantages. It has become the solution for us to two major problems. We have found it to be the best method for international data collection, since the quality of interviewing skills varies widely from country to country. Furthermore, it is best in most studies that may be used for litigation, because it eliminates one of the easiest ways a good trial attorney can dispute the findings of a study - by questioning the accuracy of the human interviewer.

IMPLEMENTING A Ci2 SYSTEM MULTIUSER LICENSE IN A STATE UNIVERSITY

by
Arthur Saltzman
Department of Marketing
California State University, San Bernardino

Introduction

During the past year my university has gone through a process of purchasing and introducing the Ci2 System for conducting computer-assisted interviewing. I will discuss this process and pay particular attention to how key decisions were made about the diffusion of this innovation at our university. Anyone wishing to use any of the techniques of forms found as appendices to this paper may do so.

First, let me describe California State University, San Bernardino. We are the youngest campus in what is probably the largest state university system in the world. There are 7,500 students enrolled, and we are still growing rapidly. Our primary focus is teaching. However, as in the case of most teaching institutions, we are increasingly asked to publish, although the pressure to publish is not as stringent as at the University of California and other research oriented institutions. I belong to the School of Business and Public Administration. We are the largest school at California State, San Bernardino. We offer a variety of undergraduate business degrees and a masters in business administration.

Now that I've introduced the environment in which this innovation was introduced, let me provide an overview of the consumer decision-making process model, which is the Engel-Blackwell-Kolat model shown in Figure 1 (page 108). I will describe the problem recognition phase; the search process; how we evaluated the available alternatives; and how the purchase decision was made. I will then digress from this consumer decision-making model to talk about the implementation process. Finally, I will return to the model to describe the outcome.

The Problem Recognition Phase

When I teach the marketing research course, I assign a live project that provides students with the opportunity to work with a real customer and produce results that will be used in some marketing decision making context. I had a serious problem with this course the last time I taught it, because we did an intercept study of more than 700 tourists and I was left at the end of the quarter with a mass of data that had been collected by the students. It then took me three months to complete the editing and coding of the information, and another month to enter the data into a computer file. This is not an unusual problem in teaching a marketing research course. We try to cram many projects into the course and frequently find we don't have time to both teach the theory of marketing research and conduct a live project. I needed to eliminate some of the bottlenecks that occur in teaching a marketing research course in this mode.

A related need is to introduce students to techniques that are currently used in the marketing research industry. All instructors have the desire to show students the currently available technology. In addition, we would like to use some of this technology in conducting our own research projects.

Search

When I first noticed the advertisement for the Ci2 computer-assisted interviewing system, I immediately sent for a demonstration disk. The demonstration convinced me that the use of a computer-assisted interviewing program would be a tremendous benefit to my class and to my research. I looked for ways to purchase the system at the university, and also searched for information on similar systems. Some of my thoughts at the time are contained in the attached memorandum (Appendix 1) that I sent to the faculty in the school of business and public administration who do survey research. This memorandum was an attempt to generate support for the university's purchase of this system.

In the memorandum I focused on a multiuser license for our campus because of the many users I foresaw. The multiuser license is still available for universities, but since the Spring of 1987, several other options might be attractive to universities that want to use the Ci2 System in stages. Given the $1500 price at the time, I needed to generate substantial support among other faculty members to find funds to purchase new software. The $1500 would cover use by the School of Business and Public Administration. For a total of $2500 all schools on our campus could have access to the software.

I was also fortunate enough to be the organizer of a panel on software piracy at the third annual American Marketing Association Workshop on Microcomputers in Marketing which took place in Pomona, California in October, 1986. I had invited Joe Curry of Sawtooth Software to participate on the panel and was able to get more details on the Ci2 System from him. Furthermore, several people at the workshop already owned and used the Ci2 System and they gave it very high recommendations. Although I considered several other similar systems briefly, there was little doubt after the Pomona conference that the Ci2 System was my first choice.

Alternative Evaluation

Although there was some competition from other brands of interviewing systems, they were quickly eliminated based on my own review of the Ci2 System and the many comments I heard from satisfied users of the Ci2 System. The next part of my alternative evaluation concerned the options available from Sawtooth Software. The Lab System was available at a relatively low price ($200). There was also a single-user license for $900 and the multiuser licenses at substantially higher prices ($1500- $6000), depending on the number of users and maximum number of questions allowed.

My needs varied. I wanted to use the Ci2 System both for research and teaching. I anticipated doing data collection with the Ci2 System, and also wanted to use it in the classroom to teach questionnaire construction. It also would be used in live cases that I assigned students in my marketing research course.

The Lab System, which allows up to 50 separate questions, was sufficient for classroom use, but was too small for any substantial research project. Also, under the old pricing system, it would have been expensive to provide enough Lab Systems for the ten-station personal computer laboratory we had at the university.

The single user license of $900 would have allowed me to do my own research, but would not have allowed any other researchers to use it on their projects. Also, it would not have allowed for use in the classroom. The third alternative was the multiuser license available at a substantially higher price. The problem was easily solved when the faculty gave me a great deal of support to purchase the new multiuser system. There was wide acceptance among the faculty of the value of purchasing the system, which finally convinced the Dean of our school to allocate funds for the purchase.

Meanwhile, Margo Metegrano of Sawtooth Software had sent me a sample system. Her letter said to test it and experiment with it for a few weeks. I was soon hooked. It became clear that even if the university didn't have enough funds to purchase the system, I would use my own funds. Fortunately, we did find funds at the university.

Purchase

Another circumstance led to a modification of our original intention to purchase the $1500 School of Business license. My live case for the winter quarter of 1987 was research for our Community Relations office on what performing arts events the local residents of San Bernardino would like to see on our campus. My students and I would prepare a questionnaire and conduct 1200 telephone interviews in the San Bernardino area. The potential usefulness of the research easily convinced our academic vice president to find another $1000 to get a university-wide license that would allow us to use the Ci2 System for this audience research project.

Implementation

In December 1987, I developed a plan for managing and controlling the Ci2 System software package on our campus. Our primary concerns were how to provide sufficient access, while not violating the terms of the multiuser license agreement. We also wanted to insure that we avoided any administrative burdens for faculty who wanted to use the system in the classroom or for research.

The details of our implementation decisions are contained in a memorandum to our Dean (Appendix 2). In summary, we allowed the computer center to be responsible for distributing the Ci2 System disk to the faculty. Those faculty who used the Ci2 System in the classroom were responsible for distributing it to their students. To facilitate the use of the Ci2 System in the classroom, I prepared a Quick-Try Tutorial that was adapted from the Ci2 System Manual "Quick Try" section with permission from Sawtooth Software, and some material that had been prepared by Bryce Johnson of Southern Oregon State University. To allow good access to the Ci2 System Manual, we placed four copies on reserve in our library. We also encouraged faculty who would be using the system intensively to spend an additional $50 to purchase their own manual.

For use in the classroom, we had a different set of problems to solve. The first had to do with the distribution and administration of many computer disks. In teaching marketing research, I conduct a large-scale survey early in the course. Although I teach the theory in the normal order that we perform in marketing research projects, I do not teach the experiential part of the course in the normal sequence. Conducting interviews early in the course allows students more insights into other phases of the process, especially questionnaire construction and problem formulation.

During the first week of class I introduced students to the questionnaire which I had previously developed for the audience research for the university. I conducted several intensive training sessions for the interviewers and in the third week of the course, we started a two-week interviewing period. We reviewed our experience during the next week, then returned to another two weeks of interviewing in order to achieve the desired sample size of 1200.

I learned several lessons from this experience. First, do not attempt to have 40 students do interviewing without adequate supervision. Fortunately, the Community Relations office provided us with a supervisor for the students. We also developed a telephone control system to keep track of the calls. Another problem we had to overcome was finding a bank of telephones with IBM-compatible computers at the same work station. We were able to find an office that was not being used after business hours, and we could start our telephone interviewing at 5:30 pm and continue it until 9:00 pm. We also interviewed on Saturdays and Sundays. The scheduling of more than 40 student interviewers was very time consuming and difficult even with the support of a supervisor.

I also wanted to teach students how to use the Ci2 System to construct questionnaires during this marketing research course. For an hour each week, I taught the use of the Ci2 System in a computer lab. This was a hands-on experience. Each student was teamed with another student at the computer. I started with the Quick-Try Tutorial, then taught them several other logic instructions and more advanced features of the Frames section of the Ci2 System. The initial start-up was difficult because half of our students had little or no personal computer experience. To help solve this, I placed several copies of the Exploring Your Personal Computer software disk on reserve in the library and recommended it to students who had no computer experience. Once we got past this problem, the students adapted well and quickly to learning the Ci2 System. The largest conceptual problem the students had was understanding that the Frames and Logic programs were linked, yet had separate functions. By stressing how these two programs work together and conducting several exercises, students became very competent and imaginative in programming their own questionnaires.

I insisted on packaging the Quick-Try Tutorial in a hard-cover, three-ring binder. I also included a plastic disk holder available through stationery stores or campus bookstores. This binder protected the students' disks. We had just one disk failure the entire quarter. This failure occurred because a student tried to save the five-dollar expense of purchasing the disk holder, tutorial and binder and instead xeroxed her friend's. With the advice of Sawtooth Software, we made the student Ci2 System disk self-booting, which required moving a few program files to include the command file on the disk. This eliminated the need for students to carry their own DOS disk in addition to the Ci2 System disk.

We were also concerned with unauthorized duplication of the software. Our approach was to take away any incentive for students to make unauthorized copies of the system by giving them easy access to legal copies of the software. Since we had a multiuser license, we could offer students the use of the Ci2 System for any university-related project. When I first gave students in my marketing research course their own Ci2 System disk, I told them they would have to return the disk upon completion of the course, but if they desired to use it on any subsequent project related to the university, I would gladly supervise the project and give them the disk. I also used this opportunity to discuss the issue of software piracy. We developed a User Agreement Form that each student signed (Appendix 3). We also made them aware that each disk had an identification frame that carried California State San Bernardino's name. With these "carrots" I also provided a "stick" by promising to fail any student caught using or making unauthorized copies of the software.

Outcome

Our experience thus far has been extremely positive. Each student successfully completed several assignments using the Ci2 System. In addition to conducting 1200 interviews for the audience research, we performed another research project on the use of personal computers by California State University students in the last five weeks of the course. This was not without great strain and stress on faculty and students. For a ten-week course, I would not recommend doing both the initial interviewing project and an entire project after that.

Students easily learned the Ci2 System. The better students were able to construct more sophisticated questionnaires. Several students became very creative with the use of color in the questionnaires and created imaginative ways of using the system. Using the system enhanced the quality of the data, decreased interviewer error, and cut down on interviewer cheating.

Several other faculty are using the Ci2 System to conduct their own research projects. In addition, the Admissions Office and Community Development Office are considering buying a CATI System to help with their fund raising.

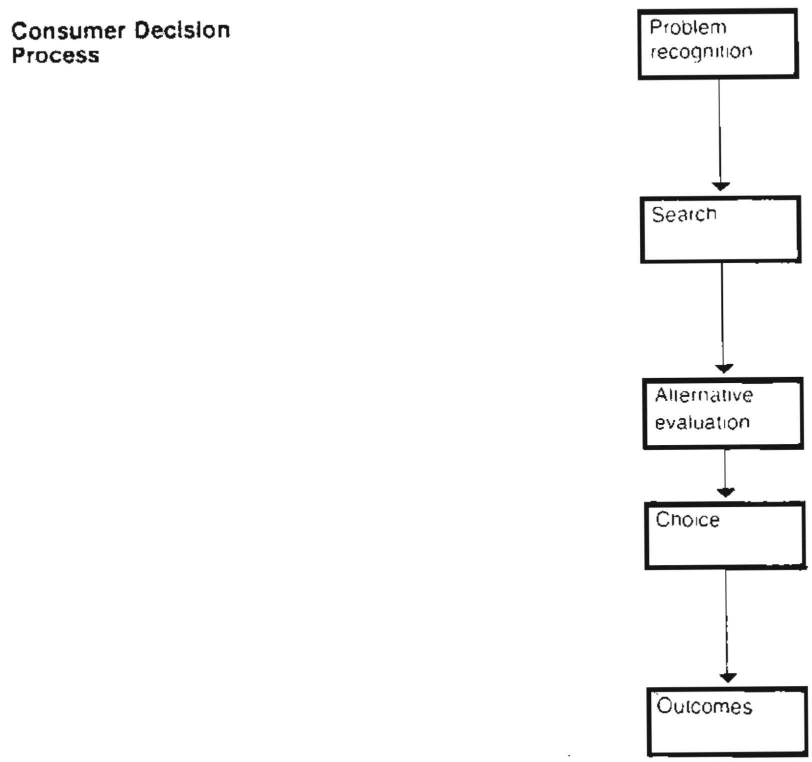
A further issue is how the use of the Ci2 System might affect students' ratings of professors. Because we are all subject to student evaluation, the effect of introducing a new technology such as the Ci2 System is a relevant concern. My evaluation by students indicated that the poorer students disliked me more and the better students give more positive feedback. This parallels the experience reported by Lawrence Dandurand of the University of Nevada, Las Vegas. My student evaluations on the average were slightly lower than usual. But the comments written on the evaluation forms indicated some excellent responses from students who appreciated the opportunity to learn a new technique in marketing research.

Summary and Conclusion

Our experience with the Ci2 System had been very positive. We have found numerous opportunities to use the Ci2 System at the university. It is easy for the students to use, and it's quite useful in conducting live research projects. However, the implementation of the Ci2 System is a nontrivial task that requires a good deal of time in the initiation phases. When implementing the system a detailed plan is necessary. The plan must delineate the implementation steps as well as who will be responsible for the many administrative tasks that will be encountered.

Figure 1

Blackwell-Kolat Model



State of California

California State College, San Bernardino
5500 State College Parkway San Bernardino California 92407

Memorandum

Date : October 8, 1986

To :

Shel Bockman	Naomi Caiden
John Chaney	Pat D'Souza
Don Lindsey	Vic Johar
Lance Masters	Janice Loutzenhiser
David Porter	Clark Molstad
Tapie Rohm	Nabil Razzouk
Mohammed Vaziri	Walt Stewart

From : Arthur Saltzman

Subject:

Collectively we conduct lots of survey research and teach many research courses. There is currently available several software packages that allow the user to use a personal computer to perform the following tasks.

1. Construct questionnaires which incorporate skip patterns, scales, multiple choice or open ended and other types of questions.

2. Conduct interviews in one of two modes.
   a. Telephone interviews where the interviewer records responses on a personal computer.
   b. Self-administered interviews where the respondent enters data into a PC and is cued by the program.

3. Generate files containing the data which are then available for analysis by statistical software packages.

Shouldn't we try to acquire some of this software for our research and teaching?

I am impressed by the Ci2 package described in the attached. I have recently talked to Sawtooth Software, the developer. They are willing to site license the Ci2 System 100 for use by all faculty and students in the School of Business and Public Administration. Price $1,500 for the System 100. The price for the System 250 would be $4,000. These prices include unlimited use for faculty and staff and more limited use (smaller lab version) by students.

Apparently these two Systems are differentiated only by the total number of responses which can be recorded for any single questionnaire. The answer to a multiple choice or yes/no question counts as one response. Answers to a ranking question generate as many responses as there are items to be ranked. Fill in the blanks or textual answers count as one response as long as they are not too large.

If we decide we like the software we could develop a "creative financing" package.

After you try the attached demo disk, let me know what you think.

CALIFORNIA STATE UNIVERSITY
DEPARTMENT OF MARKETING, MANAGEMENT SCIENCE AND INFORMATION MANAGEMENT

December 30, 1986

To: David Porter

From: Arthur Saltzman

Thank you for all your efforts towards acquiring the Ci2 System 100.

Now that this software package has has arrived I want to make some suggestions about how to manage and control it. First a quick review of the software and how it can be used.

The Ci2 System is a computer-interactive interviewing system for gathering survey information. This software package allows a researcher to compose and administer questionnaires by computer . A data file is created which is compatible with most statistical packages.

Ci2 is well documented, is menu driven, and includes:
- an editor for questionnaire construction
- a command language for designing questionnaire logic
- programs for integrating the questionnaire and logic files
- programs for creating questionnaire field disks
- programs for printing and summarizing data files

The system has been site licensed for the entire San Bernardino campus. This site license for the University will allow any student, faculty, administrator, or staff to have access to this software package for use on University business. In the classroom it will be used as a questionnaire development and data gathering tool in marketing research and survey research courses. It will also be used in survey research projects conducted at CSUSB. Its immediate use will be in a survey of audiences for CSUSB performing arts and athletic events which will be conducted by students in my Marketing Research courses (MKTG 440) during the 1987 winter quarter.

The physical product consists of five manuals and sets of diskettes containing the programs necessary to use the package. The site license allows us to make an unlimited number of copies of the software for CSUSB use. We are not allowed to duplicate the manual except for the "Quick-Try" chapter and the card which summarizes the logic instructions.

<u>Recommendations:</u>
I have several suggestions on how to manage and control this software package. They are primarily concerned with how to provide sufficient

access to the package while not violating the terms of the site license.

1. The manuals should be placed on reserve in the Library in the same manner as is currently done with other software packages. Additional copies of the manual are available from Sawtooth Software, Inc. for $50 apiece.

Sawtooth Software, Inc.
P.O. Box 3429
208 Spruce North
Ketchum, ID 83340
208/726-7772

2. The software should be administered and distributed to the faculty and staff through the Computer Center. Any faculty or staff desiring a copy of the diskettes should be required to sign a document which acknowledges the terms of the site license.

3. For use in the classroom the faculty member should require each student to sign an acknowledgement of the site license and its stipulations.

4. To allow students to use the Ci2 System for a course they need three disks and some documentation. The disks consist of a System Disk which contains the executable programs, a Development Disk which stores the questionnaires, and a Printer Files Disk for saving printer files. I recommend that the department offering the course give each student a copy of these disks to use during the course. The instructor would sign out a copy of the disks from the Computer Center and duplicate sufficient copies for students. At the end of the course these disks should be collected by the instructor. That will mean that there will be some initial cost to any department using the program. At the current price for disks purchased thrrough the University ($.55), the expense per student would be 3 x $.55 = $1.65. Thus, the initial expense to run two sections of a course per quarter with 25 students enrolled in each section would be 50 students x $1.65 = $82.50. Future offerings would only incurr expenses for lost or damaged disk which had to be replaced. In situations where there is a project being conducted for a client we could expect the client to pay for these materials. That is the case with the audience survey we will be conducting for Dean Rhymer's Office. I have requested that they pay for the 150 disks needed for teaching the students as well as the 50 disks which will be needed for collecting the data. The department could also consider recouping its expenses with a lab fee. I do not know if such a procedure has been established on our campus.

5. I do not recommend having students purchase these disks through the bookstore. This might not violate the site license but it would make it very difficult to control the use and distribution of the program. The form to be signed by students

will stipulate that the software is for class use only. If students want to use the program for another non-course related university project they should do so under the supervision of a faculty member.

6. Primary documentation for teaching students how to use the package will be the logic instructions card and "Quick-Try" chapter of the manual which will be reproduced and sold to students through the Bookstore. More detailed instructions for using the program will be contained in the manuals which will be on reserve in the Library.

The final issue I want to address is training. I am not yet an expert with the system but I would be happy to conduct a few training sessions for faculty and staff. This should probably be administered through the Computer Center. I will be contacting them about this early in the winter quarter.

DRAFT

MEMORANDUM

From: CSUSB Computer Center

To: Ci2 Users

Date: Jan. 26, 1987

Users Agreement Form.

CSUSB has acquired a survey research software package for IBM and compatable personal computers called the Ci2 System 100, which was developed by Sawtooth Software of Ketchum Idaho. With this computer-interactive interviewing system a researcher can compose and administer questionnaires. Responses to the questions are entered directly into to the computer. The interviews are recorded in a data file which is compatible with most statistical packages.

We have a Multi-User License which allows any faculty student or staff member to use this software for CSUSB purposes. However we may not permit others outside of our campus community to use it - except we can allow other agencies to use Field Disks for collection of survey data required for CSUSB uses.

The other two conditions of our Multi-User License with Sawtooth Software are first, that we label each copy of the software including field disks with the following statement:

"Copyright Sawtooth Software, Inc. All Rights Reserved."

Second, we agree not to attempt to remove the internal copyright notices or user identitication.

Faculty or staff may acquire a copy of the software necessary to use the Ci2 System 100 from the Computer Center. Students may acquire the software from faculty or staff who agree to supervise their use of the system . The student, faculty, or staff person who receives a copy of the diskette should sign below to acknowledge that they have received a copy of the Ci2 System 100 disk and that they will adhere to the conditions of the license.

One copy of this form should be maintained by the Computer Center or faculty supervisor and the other copy given to the user.

FOR FACULTY OR STAFF

Print Name ____________________

Signature ______________________

Department ____________________
or Administrative Unit

FOR STUDENTS

Name of Student ____________________

Signature of Student ________________

Faculty Super isor _________________

Course ______________________________
or Project Name

Ci2 IN THE UNIVERSITY AND TEACHING FUTURE RESEARCHERS

by
Lawrence Dandurand
Marketing Department
University of Nevada
Las Vegas, NV

First, I would like to talk about technological change. There is a great deal of technological change in the marketplace affecting marketing research. This includes computer-assisted interviewing and personal computers. It is the responsibility of marketing professors to be aware of these technological changes and to teach their students marketing practices that incorporate these changes in both marketing policy and marketing research courses.

In terms of marketplace needs, marketing research firms and corporations have an important need for someone who not only understands the theoretical aspects of marketing research, but also knows how to execute marketing research in a computer based environment.

I've had experience teaching computer-based marketing research at the University of Nevada, Las Vegas (UNLV). I have been working with the Ci2 System in particular, and computer-based research on PCs in general for three semesters. At UNLV we have three laboratories equipped with IBM PCs of different types. We also have printers, TV monitors, overhead projectors, and a PC projector, which is an IBM PC XT connected to a projector. We also have a projector pad system which is a device that sits on top of an overhead projector and is connected to a PC. It is mobile and we move it from one classroom to another. These are our hardware items.

We are receiving more support for software at UNLV. Currently in the marketing research class I work with Ci2, Microtab, Surveytab and LOTUS. One of my course objectives is to link the output of the Ci2 System to a data processing package such as Surveytab, then in turn link those results to a policy-making package. I consider LOTUS a policy-making package. I try to make that linkage between research and policy making not only in the theoretical sense, but also in the computer-based system.

My primary objective is to turn out a marketing research student who is well versed in marketing research theory. When I work with a new package, that package must help me to teach the marketing student better marketing research. Otherwise, I disregard the package because it doesn't help meet my goal of turning out a better educated marketing research student.

My particular approach, pedagogically speaking, is to ask the students to look at demos before I teach a sequence. For example, suppose we are studying questionnaire design. Before I give a lecture on questionnaire design, I ask the students to look at the demos for the Ci2 System, which is a questionnaire-administering piece of software. When they come to class, I assume that they have been exposed to the system and have a rough idea of how it works. Then I assign textbook material dealing with questionnaire design. When I lecture on the theory and the package, I weave them together and give plenty of opportunity for answering questions in a lecture-discussion approach.

Next, I demonstrate the particular software package, taking an hour of class time to introduce each new one. In the case of the Ci2 System, I work on frames and logic with the students who watch these on the monitors and projection pad. They see how the package works and gain an understanding of another dimension of the system.

After this they go into the laboratories and practice on some simple problem by themselves to become more comfortable with the system. Then I assign projects. I like field projects, because they offer the discipline of working with executives and peer members, of doing a real project, and of being responsible for real results. Students are expected to devote 90 hours of outside work to a three-credit course. About 10% of this is devoted to laboratory and field projects. I assume that the client will supply the PCs at the corporation, but the students make the field disks in the university laboratories. There are problems, of course, with coordination, student mistakes, and lost files on disks.

One always runs into those kinds of problems when working with a large number of students with complicated software packages. We have laboratory assistants and graduate assistants who attempt to train the students to prevent mistakes. When problems occur, we hope that the laboratory assistants can solve them. If they can't, then the graduate assistant is (or I am) called into the picture. After three semesters of working with packages such as Ci2, Microtab, Surveytab and LOTUS, I find it gets easier and takes less and less of my time each semester to teach more. I turn out a richer student, a more educated student. About 75% of the students are pleased with the approach. The other 25% will have to be dragged "kicking" into the 21st century.

Along with that comes the continuous concern about academic publishing. As professors, we cannot afford to spend an extraordinary amount of time helping students to learn computer-based systems, when we should be researching and publishing. There is a fine line and each professor has to make his own time judgment. I certainly am very conscious of it and take steps to insure my publications are written at the same time.

Copyright protection is another issue students need to learn about. I explain to students what the copyright is, why we have copyrights, and so on, in a non-threatening way. My laboratory assistants, graduate assistants, and students are aware that they should not make copies of any system disks and that it's strictly forbidden to take them off campus. There is no evidence that a particular package has ever been copied and used off campus. The students learning the systems probably would be more likely to recommend to a future employer that he buy his own system in order to do the project on the system they've become familiar with. Although I am very conscious of copyright and stress it to students, misuse of copyright protection has not occurred at our university.

In the future, we will have more sophisticated packages that will be integrated. We will have expert systems for making decisions in the future. Some people might think that my attempt to link the Ci2 System with Microtab, Surveytab and LOTUS is a bit too ambitious. However, students need to become accustomed to doing this as a natural course, because much more complex linkages will occur in the future. I cannot apologize for being ambitious, because the tools of marketing research are changing rapidly and the opportunities are so great, that if our students do not feel comfortable with computer-based systems, we are doing them a disservice for the future. I am very optimistic and, at the same time, very cautious about problems that might absorb my time. At UNLV the dean has been very supportive. As new, more powerful packages come out, I should have his support to purchase them.

The new academic policy for the Ci2 System is great. The license now allows us to make copies for the students in class for class projects. That helps make the Ci2 System a natural way to do marketing research. It should no longer be considered extraordinary to develop a questionnaire with the Ci2 System. Instead, I try to make it seem as natural as possible, while keeping in mind my primary objective of turning out the best-educated marketing research student possible.

PERCEPTUAL MAPPING: ITS ORIGINS, METHODS, AND PROSPECTS

by
Allan D. Shocker
Professor of Marketing
University of Washington, Seattle, WA

A useful analogy in thinking about perceptual maps is the physical maps or war boards used in the military setting. In plotting military strategy, often it is desirable to have a picture or diagram that shows the location, number, and type of enemy or friendly troop positions, weapons emplacements, and actual and manmade obstacles. Where the enemy is located; where you're located; the terrain; relative numbers; all serve to provide ideas for offense and defense. Military analogies have been used quite frequently in talking about managerial issues in the context of marketing warfare. Such ideas by Phil Kotler, Trout, Reis and others have received quite a bit of prominence. Maps fit in very well, at least potentially, with this view of the world. Maps enable planners to communicate better with others, to see the big picture, to see the role they're expected to play in that picture, and to have some idea of the magnitude of the task that lies before them. It's useful to think of perceptual maps and managerial maps in somewhat the same way. Yet there are significant differences between them.

Historically, perceptual product mapping has many variants, not just one by any means. Despite the fact that you'll hear about one particular approach to mapping produced by Sawtooth Software, it's simply one of many different methods and techniques that exist in this area. There are several major classes of techniques that I would like to familiarize you with briefly.

A perceptual map is intended to be a <u>model</u> of how a market sees a set of products or competing products in some memory-like or cognitive sense. In reality, because multiple people are involved and they may see things quite differently, the map is at best a crude approximation of what might take place in anyone's memory. Indeed, some would argue it is not a very good representation at all.

There are three general types of methods: Fully metric methods were historically the first to be developed. They are based largely upon factor analysis and discriminant analysis and follow from theorems by Young and Householder. Their task was to find the appropriate number of dimensions and to represent a set of products or objects by a spatial configuration. Fully nonmetric methods came historically next and trace back to the work of Clyde Coombs in the early 1950's. Such methods allowed ordinal input judgments to be analyzed, except that these methods also produced only ordinal output. They positioned products on a set of attributes in rank ordinal form and did not provide the configuration characteristic of the fully metric methods.

In 1962 Roger Shepard published a now classic work that tried to abstract the better features of both the metric and the nonmetric methods in what he called "Nonmetric Multidimensional Scaling." Here the task is to take the given ordinal input and use it to find a cardinal (ratio or interval) scaled configuration of products in this multidimensional space. Nonmetric methods, therefore, can be viewed as a way of transforming ordinal measures into higher order or cardinal ones. These methods were developed by Joseph Kruskal, who produced a procedure for implementing Shepard's ideas. As a consequence, several programs have been developed to implement nonmetric approaches.

As noted, metric methods are exemplified most commonly by factor analysis and multiple discriminant analysis based procedures. They are called composition models, in part, because they take aggregated data in the form of individual perceptions of products on a set of prespecified attribute scales and group them into a composite whole. These methods produce a perceptual configuration in a reduced number of dimensions from that with which they originally started.

Nonmetric methods, as exemplified by nonmetric multidimensional scaling, are decompositional methods. They ask people for summary judgments. Then they decompose them into some understandable representation of those products or objects on a set of dimensions. Those dimensions, unfortunately, need to be interpreted (as also do the metric methods), although the basis for that interpretation differs in the two cases.

It's useful to think of these two classes of methods, metric and nonmetric, as complements rather than substitutes. The advantage of the nonmetric methods, particularly, is that they are less context laden. They require people to judge the perceived similarity of a set of products without telling them the criteria they are to use to make their judgments. As a consequence, the methods are more likely to reveal information about the attributes that underlie such judgments. In fact, they were originally intended by the psychometricians who developed them for that purpose. They give you an opportunity to peer into the consumer's mind by allowing the consumer to impose his own structure, and this may even suggest characteristics or attributes of products that might otherwise not have come to mind.

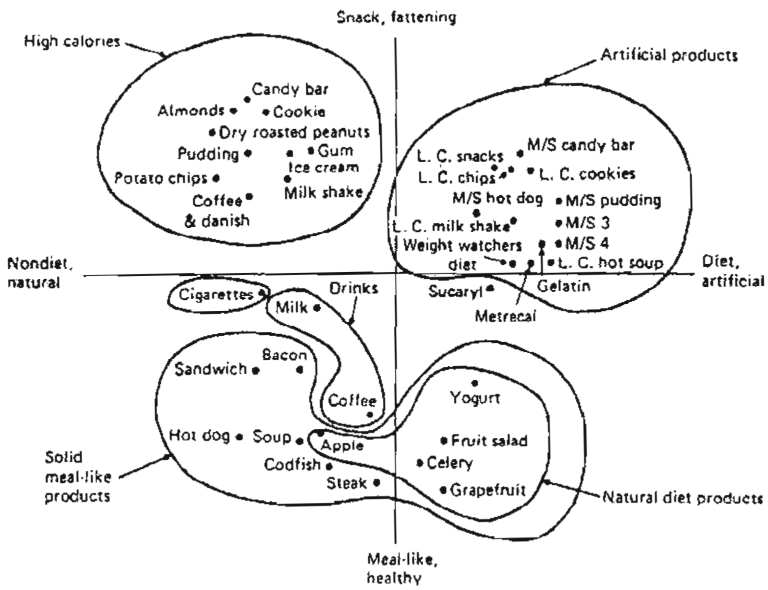
The metric methods require a prespecified set of attributes and a scale explicitly developed to measure where products are on them. As input, they ask consumers to rate or evaluate products on those prespecified product dimensions. This places a great burden on you to correctly develop a set of attributes and ways of scaling products on those attributes before the analysis begins. If the attribute set is incomplete in some fashion, or if there are problems in getting people to encode their perceptions of products on the relevant set of attributes, there can be problems with subsequent analysis. For these reasons, the two methods should be thought of as complementary.

Both methods have other advantages and limitations. Ordinal measurement is easier for respondents to make. They simply have to identify, in the case of nonmetric scaling, whether one pair of products is more similar than another pair. They are not asked to quantify the degree of similarity. If preference data are used as input, they have to indicate whether they like Product A better than Product B. Such measurement is easier for customers to make and, therefore, tends to be more reliable. When you ask people to rate things, you're asking them to quantify their judgments on a scale that they've never seen before, that a researcher has provided, and as a consequence, they may find that task difficult. They may, in fact, change their recorded perceptions from time to time and become less reliable in their judgments. On the other hand, what you lose using nonmetric methods is statistical power. There are a variety of statistical tests available for the metric methods that are absent, or at least diminished in power, in the case of nonmetric methods. These are a couple of the major differences between these two categories of technique.

Figure 1

TWO-DIMENSIONAL PERCEPTUAL CONFIGURATION OF 27 FOOD PRODUCTS AND 13 NEW DIET CONCEPTS

Source: Reprinted from "Product Positioning: An Application of Multidimensional Scaling," by Y. Wind and P.J. Robinson, in *Attitude Research in Transition*, R.I. Haley (ed.), published by the American Marketing Association, 1972.



TWO-DIMENSIONAL PERCEPTUAL CONFIGURATION OF 40 PRODUCTS AND THEIR "PERCEIVED" AND OBJECTIVE ATTRIBUTES

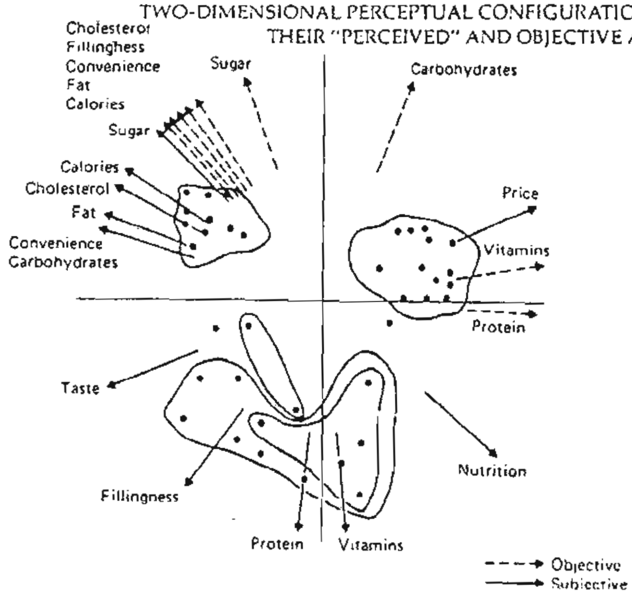


Figure 1 is an example of a map that has been created by nonmetric multidimensional scaling of a number of snack-type products. The lines drawn around some items represent the author's attempts to cluster different snack products. Both metric and nonmetric techniques allow you to correlate a set of prespecified attributes (that have been developed independently in the case of nonmetric methods) to help in interpreting the underlying dimensions. You thus have the perceived locations of the products on these prespecified attributes, as well as the estimated product proximities to guide you in that interpretation. In this example the authors hypothesized a "diet/artificial" "non-diet/natural" horizontal dimension and a "snack/fattening" "meal-like/healthy" vertical dimension that arises out of that interpretation. Both objective attributes and subjectively scaled ones can be used to aid this interpretive process, as this example illustrates.

Among compositional methods, the principal ones used for mapping have been factor or principal component analysis and discriminant analysis. Consider some of the distinctions between these two approaches. Figure 2 shows the discriminant and factor analysis solutions to an identical set of data. The authors who proposed this example, Joel Huber and Morris Holbrook (Journal of Marketing Research, November, 1979), have created a situation in which you have one attribute, in this case "fast/slow," which is a relatively objective attribute. (The presumption is that people can reliably make judgments on such a characteristic.) On the other hand, the second characteristic measures what people like, i.e., "good/bad," and is a more subjective characteristic on which people are expected to hold greater differences in opinion. Huber and Holbrook postulated a set of individual perceptions of three different products on these two scales, and fitted discriminant functions and principal component functions to that same data. Multiple discriminant analysis takes account of individual perceptions of each product and tries to find a function along which people will discriminate between three products: A, B and C. In Figure 2, A, B and C represent the three products; 1, 2, 3, 4 and 5 represent five different people's perceptions of where those products might lie in these underlying dimensions. In this case, discriminant analysis tries to find a dimension along which the variance of the projections of people's perceptions of the products will be maximally distinguished with minimal disagreement. More technically, discriminant analysis tries to maximize the differences between the centroids of these perceptions of the three different products, while at the same time, minimizing the variance of individual perceptions about any single product. Consequently, its objective is to maximize a ratio of between to within variance.

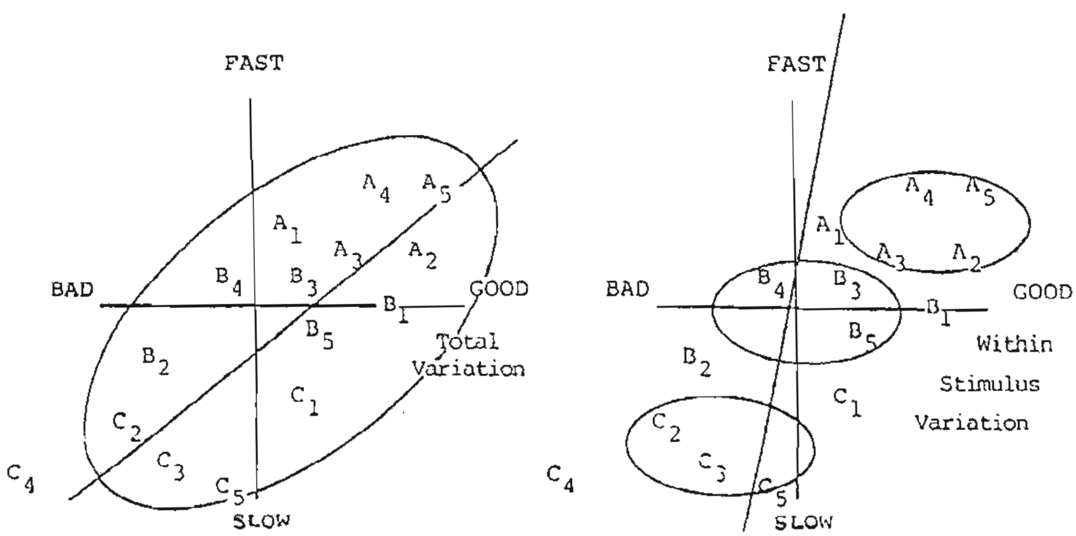
Factor analysis doesn't distinguish between the products initially. Instead, it examines the relationship that exists between different attributes over a combination of products and people, so that it treats the entire data set as a whole and looks at the relationship that exists between the underlying attributes. As a consequence, it produces somewhat different results shown in panel 2 of Figure 2.

Figure 2

ILLUSTRATION OF THE DIFFERENT ORIENTATIONS OF THE FIRST PRINCIPAL COMPONENT AND THE FIRST DISCRIMINANT FUNCTION

(evaluation of stimuli A, B, and C for 5 respondents)

Panel 1	*Panel 2*
The First Principal Component	The First Discriminant Function



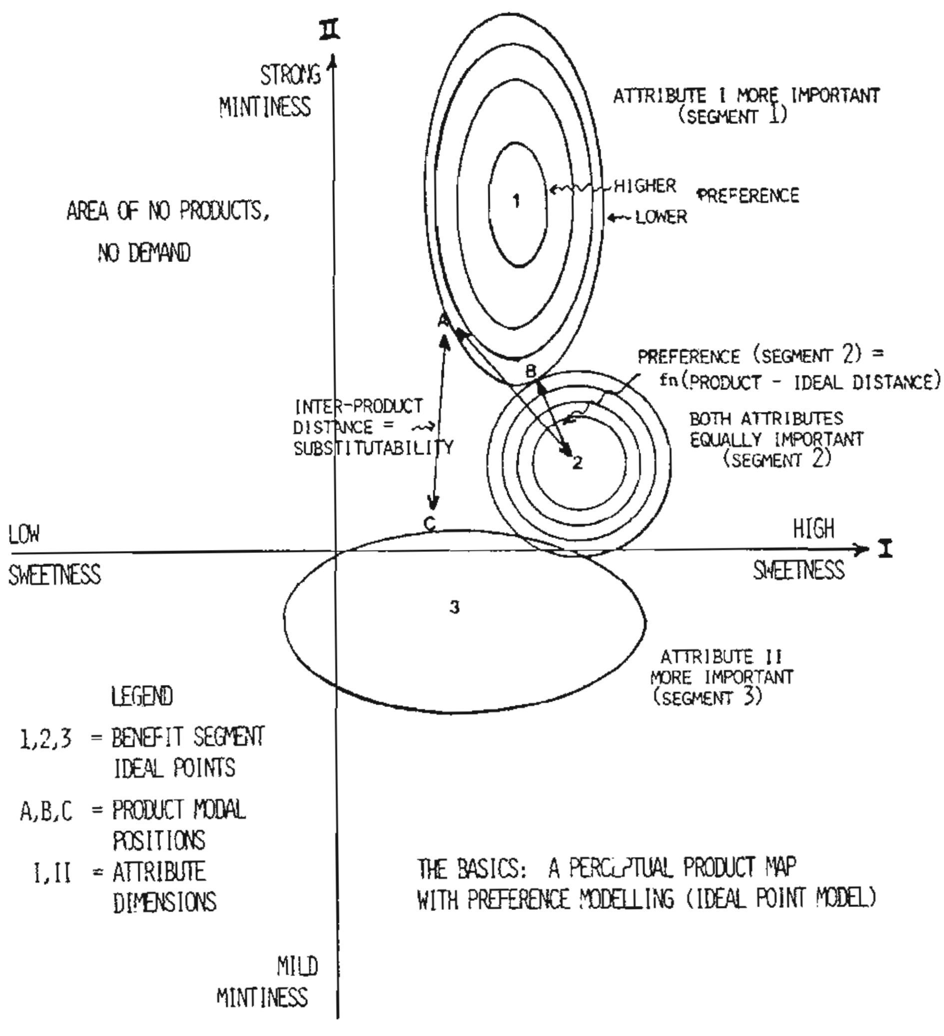
This diagram was meant to argue that. The discriminant function emphasizes its proximity or closeness to each original attribute by the angle it makes with each original attribute. In this case, there's a small angle involved, which implies this discriminant function is more closely aligned with the "slow/fast" than the "good/bad" dimension. According to Huber and Holbrook, discriminant analysis will be most influenced by more objective dimensions in which there is a relatively high level of respondent agreement. On the other hand, factor analysis has a somewhat different orientation. It's more of a 45-degree angle (panel 1 of Figure 2) in this case and oriented toward both attributes approximately equivalently. It is more likely to reveal the presence of these two underlying factors than is the discriminant function.

Hauser and Koppelman, who have published results on different comparative approacnes to perceptual mapping, have argued that factor analysis is superior to discriminant analysis in terms of the theoretical justification (Journal of Marketing Research, November, 1979). They found more satisfaction in this attempt to understand the correlations that exist between attributes, rather than to focus on distinguishing the products. They used holdout samples of people's judgments. They found that multi-attribute models of customer decision-making based or premised on a factor analytic derived space tended to have higher predictability. They also found in their empirical results that the dimensions that were extracted through this method had higher managerial interpretability than those that came out of discriminant analysis.

In addition, they also noted that both methods, because they rely on widely-available software, were equivalent in terms of ease of use, but dominated multidimensional scaling, which relied on less readily available, special-purpose software. Discriminant analysis tends to extract dimensions which more nearly resemble those attributes about which most respondents agree in their product ratings, i.e., are more objective in reflecting homogeneous perceptions across respondents. They have the desirable characteristic that they produce attributes that are more actionable, so that you as the manager can better understand what to do as a consequence of differences in perceptions on those attributes. Factor analysis, on the other hand, often provides solutions more sensitive to the "meaning" of attributes, because it doesn't look at the product per se but tries to understand dimensions that the attributes in question have in common.

Perceptual maps not only can capture competitive relationships between products (or people's perceptions of products), but they also are able to capture some aspects of customers' choices among them. In other words, not only do they tell you something about how different products are perceived, but also where or what customers desire, and which customers are being satisfied by different product alternatives.

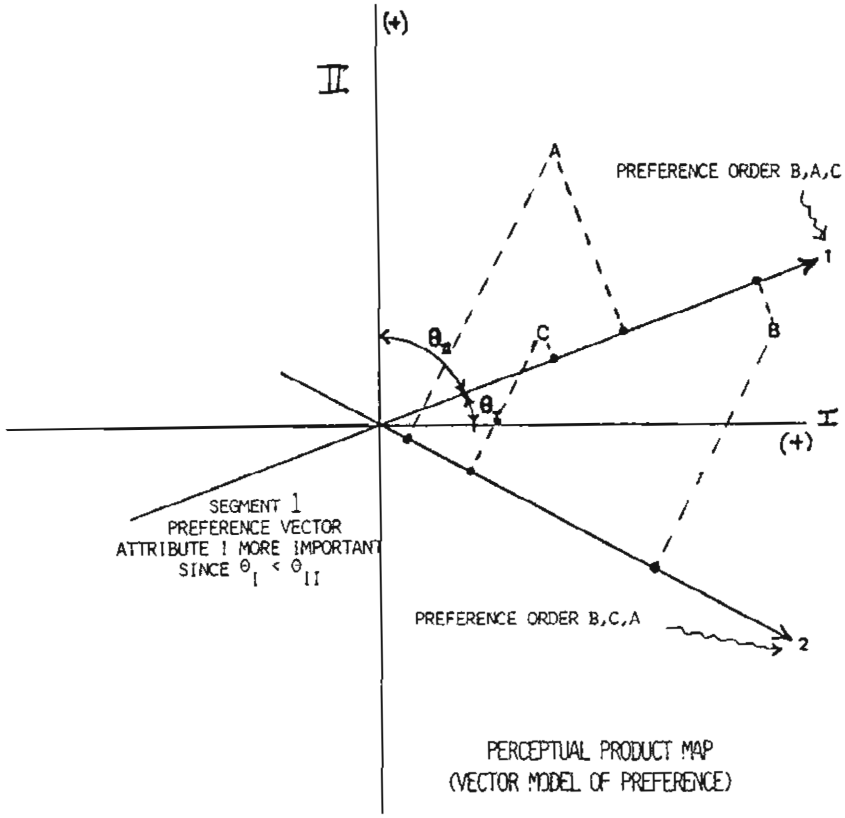
Figure 3



There are two major models that have been used with the metric methods for generating perceptual space: the ideal point model and the vector model. These imply different things about the way customers are presumed to make decisions. Figure 3 shows the ideal point model in a perceptual space. This is a fictitious example to illustrate the models themselves. Each customer or customer segment, represented by numbers, is presumed to have some ideal level of the underlying attributes. In this example, a most-preferred level is presumed with two attributes: mintiness and sweetness. The presumption is that you can have either too much or too little sweetness or mintiness and each is equally bad relative to a "just right" level. Customer One, in particular, prefers a relatively high level of mintiness and a relatively moderate level of sweetness. Customer Two prefers greater sweetness and substantially lower mintiness. If you move to either side of the ideal level, preference diminishes on the part of each of these customers. Customer One is represented by a series of isopreference contours that are elliptical in shape, indicating that Customer One does not weight the two dimensions equivalently. The relative importance of sweetness and mintiness to each customer can be represented by the degree of elongation, assuming that sweetness and mintiness are measured in comparable units. The shorter dimension indicates the dimension more important to the individual in the sense that an equivalent movement along the more important scale would have a larger effect on preference. Customer One values sweetness more than mintiness. Customer Two, his isopreference or lines of equal preference being circular, indicates that he values the two attributes equivalently. Customer Three values mintiness more than sweetness.

The vector model, as its name implies, represents customers by directional vectors. The model, as shown in Figure 4, assumes that individuals want as much (or as little) of the attributes as they can get, and they value products which give them more (or less). Customer One is depicted as valuing the products in the rank order of their projections onto his vector, i.e., B, A, C. Customer Two, or segment two, has a different rank order - B, C, A. As noted, the presumption is that individuals want as much or as little, and choose products accordingly. Again, the orientation of the vector allows one product attribute to be more (or less) important than another. The two models imply something different, although mathematically the vector model is a limiting case of the ideal point model.

Figure 4



Both metric and nonmetric methods again tend to be used to conceptualize the two ways of modeling customer judgment within the confines of the perceptual map. Metric methods are exemplified by multiple regression and logit analysis (an econometric technique that analyzes people's choices as opposed to their preferences). Nonmetric methods also exist. LINMAP, MONANOVA and PREFMAP are such alternative methods for developing ideal point or vector models or combinations of these. The empirical studies that have compared these different results tended to indicate that multiple regression can produce roughly equivalent predictive results to those obtained by LINMAP and MONANOVA. However, LINMAP permits greater modeling flexibility because it allows the use of different models of decision-making for different attributes. If your understanding of the perceptual map dictates that some attributes are better modeled as vector and some as an ideal point, you can mix the two in the same overall decision framework. You can model customers as making some of their attribute choices in vector terms, while other attributes can be modeled as if they had an ideal point. It also allows the combination of so-called partworth or conjoint type models to be mixed with ones that assume continuous attributes.

The other advantage that LINMAP affords relative to metric methods is that it allows you to impose a priori constraints on attribute weights. If you have a prior notion that preference ought to increase with attribute level, or that certain attributes are more important than others, you can find the solution most consistent with the data that satisfies that particular constraint. If logic implies that people ought to have lesser preference for higher prices, you can impose that upon the solution so that it best fits those constraints. The reason for doing that is the data themselves. Empirical data can sometimes be misleading. If you're totally driven by what the data tell you, as opposed to having any kind of logic, you can be misled in many ways because of measurement problems. Having the ability to impose some theoretical understanding may afford better predictive insights than otherwise. Those are the major advantages of LINMAP vis-a-vis some of the other methods that I've alluded to.

The reason for being concerned about which model to use, ideal point vs. vector, has to do with the nature of the attributes involved. In 1981, Jim Meyers and I published a paper in which we speculated and offered some empirical support for the notion that ideal point models will prove more useful or appropriate when the underlying attributes are interpretable in terms of physical or pseudophysical product characteristics. Pseudophysical product characteristics are not uniquely linked to a specific physical ingredient, but have some of the same properties. Managers often know how to change that attribute through manipulation of specific physical features. Examples of pseudophysical attributes would be "spiciness" and "softness," which have some physical analogies, but are not linked specifically to the variation of a single ingredient. When your attributes developed on your perceptual maps are physical or pseudophysical in nature or can be construed to be so, then an ideal-point model should work reasonably well, and there should be some limit beyond which people's preferences drop off. The vector model is more appropriate in cases where the underlying space is interpretable in terms of benefits or costs. Presumably, people want as much (or as little, in the case of cost) as they can obtain. Therefore, you might prefer to model their decision making in a vector format.

All of this has given rise to a fair number of applications for mapping and mapping methods. Figure 5 shows a laundry list largely linked to academic published literature. Practitioners may know of still other things that they have done with perceptual mapping. To whet your appetite for the potentialities that exist with this method, I'll allude to some of the kinds of things that one can do with such maps.

Figure 5

DETERMINATION OF MARKET STRUCTURE

MARKET DEFINITION	STEFFLRE 1968; DAY, SHOCKER, & SRIVASTAVA 1979
BASES OF COMPETITION	GREEN 1975; JACKSON & SHAPIRO 1979; PESSEMIER 1982
DETERMINE BENEFIT SEGMENTS	GREEN & CARMONE 1970; PESSEMIER & ROOT 1975; GINTER & PESSEMIER '78
PRODUCT/SERVICE BUNDLES	SRIVASTAVA, LEONE, & SHOCKER 1981 SRIVASTAVA, ALPERT, SHOCKER '84
PRODUCT IMAGE/FAMILIARITY	GOLBY 1968; MACKAY & ZINNES 1981; PESSEMIER 1982
CANNIBALIZATION POTENTIAL	LEHMANN 1971; ASSAEL 1971; KALWANI & MORRISON 1977

SIMULATION OF MARKET BEHAVIOR

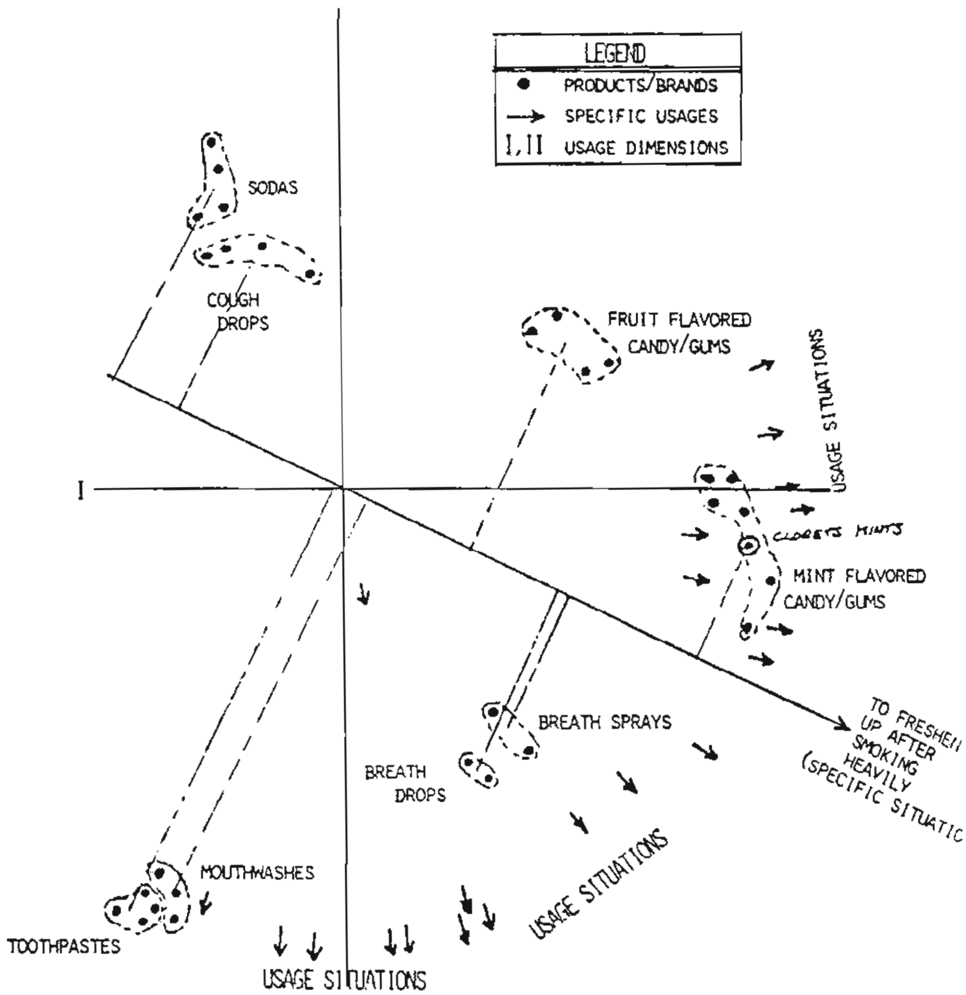
POSITIONING/REPOSITIONING	JOHNSON 1971; WIND 1973; ARABIE, CARROLL, DESARBO & WIND 1981
EVALUATE NEW PRODUCT	WIND 1973; SHOCKER & SRINIVASAN 1974; URBAN 1975
PRODUCT DELETION	PESSEMIER 1982
OPTIMAL NEW PRODUCTS	GREEN, CARROLL & GOLDBERG 1981; SUDHARSHAN, MAY & SHOCKER 1987
GUIDANCE TO R&D/BOUNDARY OF FEASIBILITY	URBAN & HAUSER 1980
MARKET ENTRY/ MERGER-ACQUISITION MARKET POSITION	URBAN & HAUSER 1980; WIND 1982
ANTICIPATE COMPETITIVE ACTION/DEFENSIVE STRATEGY	HAUSER & SHUGAN 1983; HAUSER & GASKIN 1984

DYNAMIC MARKET DEVELOPMENT AND EVOLUTION

AID COORDINATION/ COMMUNICATION	
MONITOR/CONTROL MARKETING ACTIVITIES	VAVRA 1972; URBAN 1975; URBAN & HAUSER 1980
TRACK MARKET EVOLUTION OVER LIFE CYCLE	GREEN & CARMONE 1970; GREEN 1975; PESSEMIER 1982
FACILITATE ORGANIZATIONAL LEARNING	SHOCKER & SRINIVASAN 1979; PESSEMIER 1982

Several people, including myself, have been associated with the use of maps in defining or helping define markets. One of the nice properties of maps is that you can operate with them on almost every level. You can look at competition among product types, as well as competition among individual brands within a product type, and between entirely different product categories. Figure 6 shows a map that resulted from some early work in examining the breath freshener market, and particularly, trying to discover what products competed with Clorets Mints. We had a large number of different product types. While experimenting with this method, we added some products unrelated to breath freshening, such as soda and coughdrops to see what the methodology would do. Fruit-flavored candies and gums, mint flavored candies and gums, breath sprays, drops, toothpastes and mouthwashes were other product categories. Our approach was to regard as competitive those products which were judged to be appropriate to a similar set of situations.

Figure 6



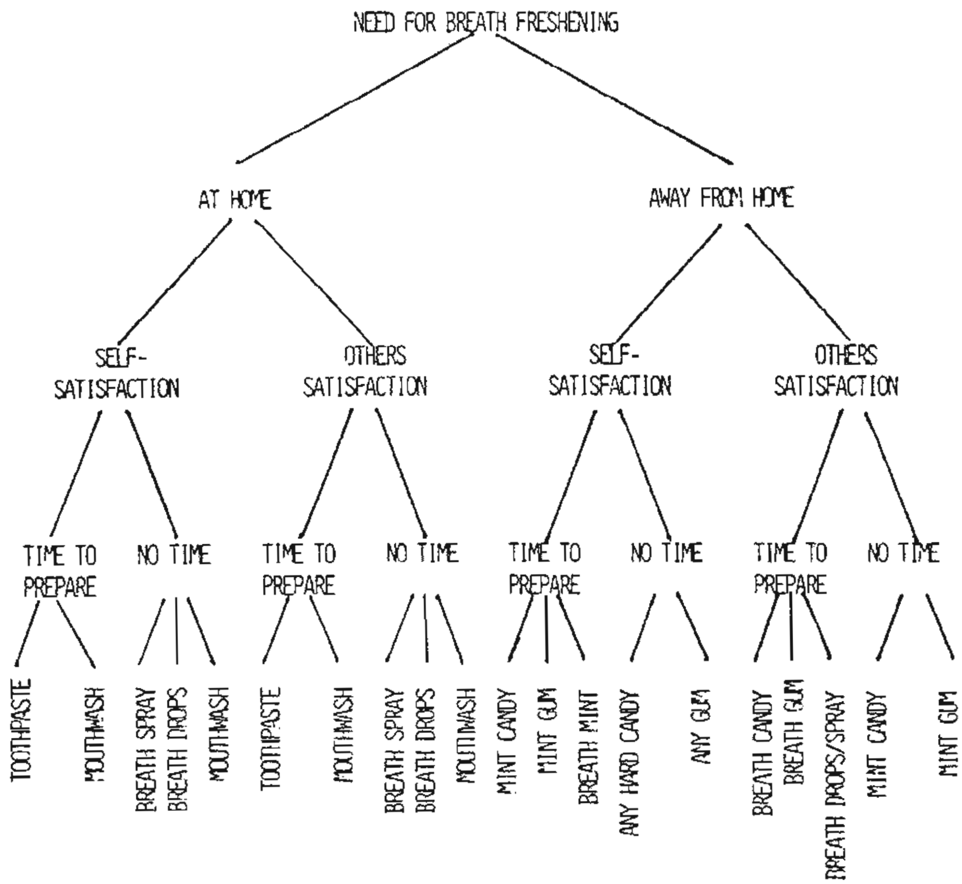
MARKET DEFINITION USING PERCEPTUAL MAPPING
(PRODUCT - USAGE ASSOCIATIONS)

We looked at what people did with the products and the circumstances under which they might use different products, the premise being that the products compete to the extent that they are used in similar ways. We plotted a set of situations or usages to help us visualize the product-usage association. I've indicated in Figure 6 how one might interpret that product-use association. Products that project heavily toward the direction of the usage situation arrow are very likely to be used in those situations. Products that project toward the tail of the arrow are less likely to be used in such situations. Since all the situations are in the lower right-hand quadrant of this map, the sodas and cough drops tend not to be used as breath freshening agents. Maps allow you to see where things plot and to draw clusters or clumps around them in order to understand what specific brands compete with each other.

One surprise of this study was that for mint-flavored candies and gums, flavor tended to be more important than form (i.e., candy vs. gum) in determining which products would cluster for breath freshening types of applications. At the time, Closeup toothpaste had just been introduced and it purported to be between a toothpaste and mouthwash. Nonetheless, Closeup was very clearly in the toothpaste category, rather than in the mouthwash category in terms of people's perceptions. Closeup wasn't the hybrid product that had been anticipated. By starting with a broad array of products, you can look for clustering to give some idea of who the competitors might be to your brand and to other brands.

This particular study also enabled us to identify a set of characteristics that the usage situations had in common and to create a taxonomy of usages that would help us better understand why products were used when they were used. Out of that comes some understanding of why things compete with each other. Figure 7 is an attempt to group the products together in terms of their appropriateness to different kinds of usages described in terms of the underlying dimensions. In the breath freshener study we found that being at home vs. being in the office or outside the home played a role in which product might be chosen. The distinction between whether you were trying to freshen your own breath, as opposed to your concerns about how others perceive your breath freshness, tended to be important, as was the anticipation of the need for a breath freshener. If something came on you suddenly, you might choose a different set of products than otherwise. This method provides some clues as to why the structure of a market is as it is.

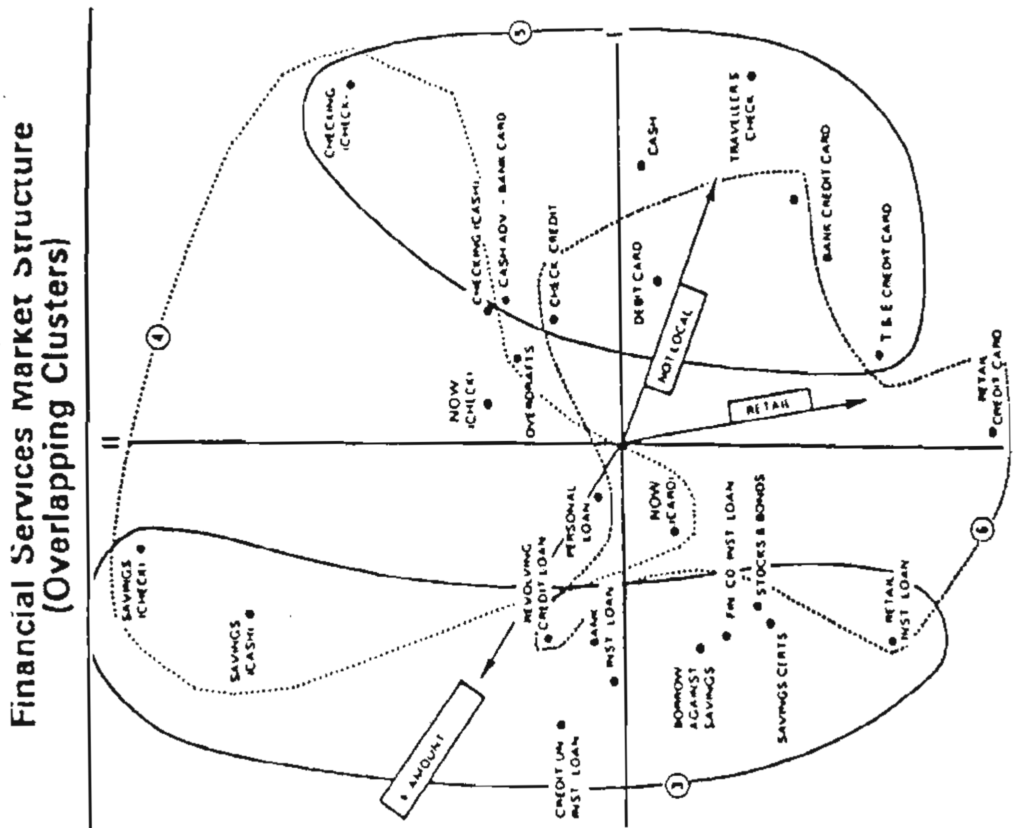
Figure 7



USAGE-DEFINED SUBMARKETS BASED UPON OVERLAPPING CLUSTERING

We did another study with a larger database that examined the competition between financial services. Attempts were made using a method of overlapping clustering to group these products into competing subsets as shown in Figure 8. The lower table attempts to explain the clusters in terms of the kinds of situations that represent each and then to associate the different products that were strongly related to each of the clusters. One of the advantages of this approach is that it helps to understand what products might usefully be put together into a bundle of products with a joint price. The sponsors of this study were looking for some criterion to carry out that bundling of products that they would offer to consumers as a package at a discounted price. The criterion we used was to find a small set of products that satisfied a large set of different usages for consumers, and to remove products that were used in similar ways from that bundle.

Figure 8



ADCLUS Groupings for 24 Financial Services

Order of Entry ⇒	1	2	3	4	5	6
Rank by Weight ⇒	1	4	2	5	3	6
Cluster Weight ⇒	64	43	56	35	50	31
Cash on hand		X			X	
Checking (cash)		X		X		
Checking (check)		X		X	X	
Debit card		X			X	X
Savings (cash)	X		X	X		
Savings (check)	X		X	X		
NOW account (check)	X			X		
NOW account (cash card)	X			X		X
Savings certificate	X		X			
Stocks and bonds	X		X			
Borrow against savings	X		X			
Traveller's check		X			X	
Cash advance–bank card		X		X	X	
Overdraft–checking		X		X		
Check credit–bank card		X			X	X
Bank credit card		X			X	X
T&E credit card		X			X	
Retail credit card		X				X
Personal loan–relative	X			X		
Bank installment loan	X		X			
Finance co. installment loan	X		X			
Credit union installment loan	X		X			
Retail installment loan		X	X			X
Revolving credit loan	X		X			X

Interpretation of Clusters (in order of entry)

(1) Products likely to be used when large amounts are required (local use).
(2) Products likely to be used when low/medium amounts are required in retail settings.
(3) Products likely to be used when large amounts are required in retail settings (local use).
(4) Products likely to be used when low/medium amounts are required in nonretail settings (local use).
(5) Products likely to be used when low/medium amounts are required in out-of-town settings.
(6) Products likely to be used when medium amounts are required in retail settings while out of town.

A worry in introducing new products is cannibalization, and maps can be used to discover what effect a new product introduced to the market will have on a company's existing products. An intriguing combination of maps with a hierarchical structure of the coffee market appears in Figure 9. This approach was popularized by Glen Urban and John Hauser. Figure 9 comes from their textbook. They looked at how the coffee market was partitioned: instant vs. ground; caffeinated vs. decaffeinated; regular vs. freeze-dried. Then they developed maps of products that competed within each partition. Figure 9 shows the brands in each partition and whether a firm has multiple brands in a single partition which would indicate a greater possibility of cannibalization. The same could be done for previous kinds of maps with clusters to see whether new products and existing products would be grouped together in a single cluster.

Figure 9

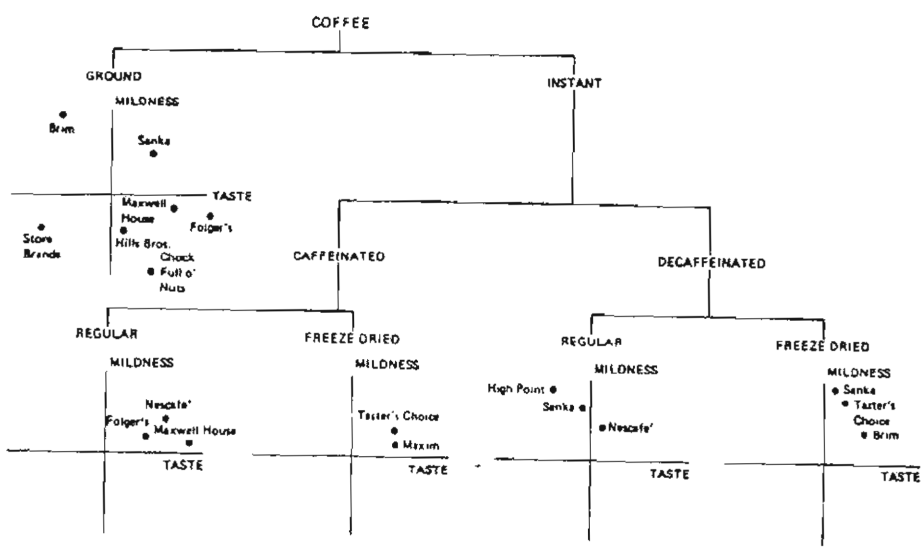
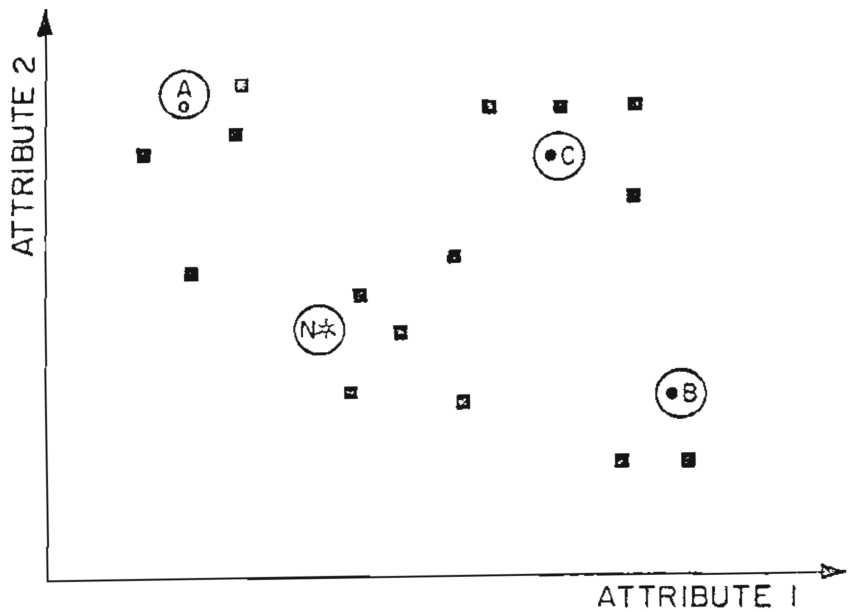


Figure 5.7. *Hierarchical Definition of Coffee Market (Urban, Johnson, and Brudnick, 1979)*

Maps have also been proposed as a method for positioning products. Figure 10 is a rendering of one such application where the circled objects A, B and C are existing brands. The dots or squares indicate locations of demand as measured by locations of ideal points, and "N" represents a proposed positioning for a new product. Some customers want products in this vicinity, and they are not being adequately served by the existing products. This suggests the opportunity for a new product or a repositioning of an old one.

Figure 10
PRODUCT POSITIONING



LOCATION OF EXISTING BRANDS

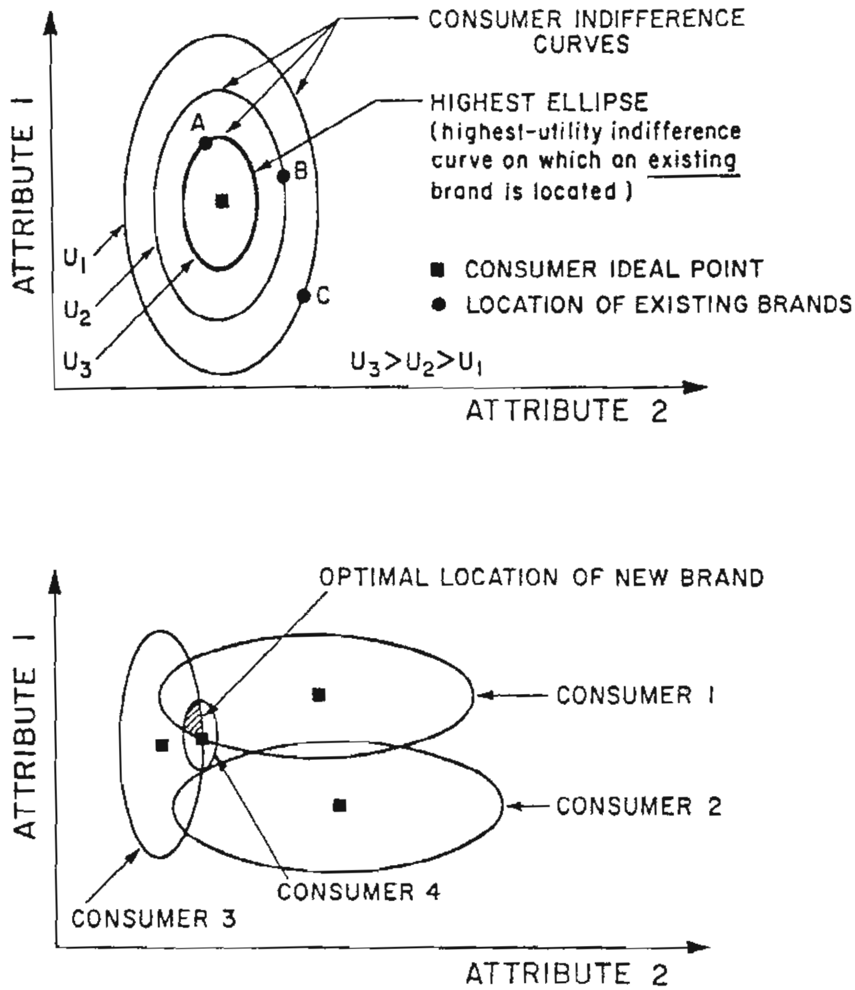
LOCATION OF CONSUMER IDEAL POINTS

POTENTIAL LOCATION OF A NEW BRAND (N)

Maps have been suggested as useful in helping managers find promising new product ideas, and several software producers have optimization methodologies for this purpose, but it's not clear how useful such methodologies are. A great deal more research is needed to determine how useful these are. If you have a group of consumers who have a set of isopreference contours, and you have some way of discovering how they will choose brands, the presumption is that they will choose whichever brand is closer to their ideal point. Then one can map the closeness of the set of existing brands to each of these customers to find a location for a new brand that will simultaneously be closer to the greatest number of customers. Figure 11 is a geometric analogy of how that might be carried out. Any product located in the shaded region will simultaneously capture the demand of three segments and, therefore, should be a product that is worth introducing.

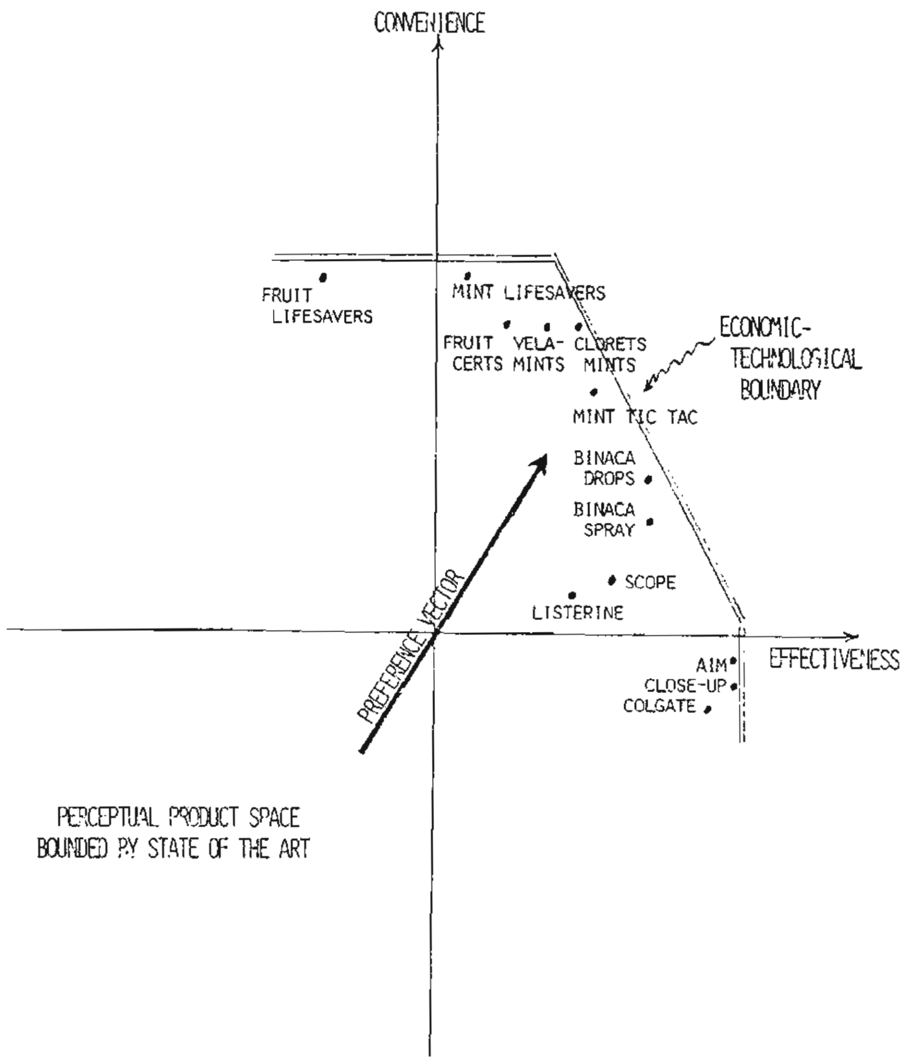
Figure 11

CONCEPTUAL EXPOSITION
OF
NEW PRODUCT MODEL



An intriguing use has been suggested, again by Urban and Hauser, of using maps to guide research and development spending. Figure 12 shows how we plotted a set of existing products in perceptual space. The boundary around the set of existing alternatives represents a region that may not be penetrated for economic or technological reasons. Since we have some ability to model people's preferences or choices, we have some ability to predict how people would respond if a product could be located beyond that boundary. In other words, if R & D could produce a product that was simultaneously more effective and more convenient, what would it be worth and what kind of demand might it capture? That insight could be useful in helping firms prudently establish R & D budgets.

Figure 12



As useful as maps are, they're not analogous to the physical kinds of maps that we associate with military applications. For one, maps don't have a physical reality with which the map can be compared. With a physical map, you have the terrain itself, and you can see how well the map models that terrain. You can get aerial photos to verify the spotting of positions. Few such standards exist in the perceptual environment and, therefore, it's important that any inferences drawn from mapping be validated with other independent research, rather than be taken as truth. Perceptual maps have the further disadvantage that while they represent competitive and demand factors, they ignore the costs of creating new products or repositioning existing ones, which might be important in determining profitability. Some highly preferred products may be too costly to build. People who work with maps have seen this phenomenon.

There is considerable disagreement in the literature regarding the appropriate measure of competitiveness to use in modeling applications. Nonmetric multidimensional scaling has used perceived similarities, or similarity of preference judgments. More recently, some work has developed which uses behavioral data and scanner data to develop maps. Others have proposed the discriminant analysis, vector analytic approach which tends to rely on people's judgements on prespecified attributes and similarity of perceptions as a proxy. Different methods exist, and it's not necessarily clear which are best or even better.

The software Rich Johnson has developed uses discriminant analysis as its underlying approach. Factor analysis has thus far been the more widely used approach to developing maps in the academic literature. The existence of the Sawtooth Software system should result in the greater use of discriminant analysis. That is a two-edged sword. It's very easy to use Rich Johnson's program and, as a consequence of its ease of use, people will use it without necessarily understanding the basis behind it. The danger in making things easy is that techniques tend to be misused. Most mapping methods fail to guarantee the actionability of the underlying dimensions. They rely on interpretation which may or may not prove insightful. You usually don't know what products will fill gaps in perceptual space and, therefore, it's useful to introduce a product and then map again and see whether or not you've been able to realize the attained position.

Different usage contexts, different products, and different degrees of familiarity affect mapping. New products tend to be less familiar to people than older products and, consequently, their perceptions may be more heterogenous. People tend to use different decision rules even though you may model them as though they use the same rules. A map is at best a snapshot of what exists in a moment in time. It takes a great deal of skill and insight to be able to take a single map and project how the marketplace will change as a consequence of changes that competitors introduce. Those who want to learn more about the foibles of mapping should read a very good article by Dillon, Frederick, and Tangpanichdee that appeared in the June 1985 issue of The Journal of Consumer Research.

ADAPTIVE PERCEPTUAL MAPPING

by
Richard M. Johnson
Sawtooth Software, Inc.
Ketchum, ID

The usefulness of perceptual maps in marketing has been recognized for about 25 years. Techniques for making maps differ; there is no single, best method, but there is a way that I've found very useful. It's a method I became aware of in 1968 and published in a 1971 issue of the Journal of Marketing Research. That method, with some modern-day improvements, lies at the heart of our Sawtooth Software product, the Adaptive Perceptual Mapping (APM) System.

I'm a little red-faced about giving a commercial message this morning. On the other hand, the APM System is a tool that researchers should know about even if they never use it, because it embodies several widely used, generally useful, and relatively fool-proof ways of doing things. I've been using these methods since the 1960's, as have dozens of my colleagues. I have never heard of an unpleasant surprise resulting from their use. They appear to be well behaved, and produce useful results time after time.

Those of us present today occupy a broad range of background and familiarity with perceptual mapping. I will keep this presentation at a non-technical level, but additional details can be found in a paper, "Adaptive Perceptual Mapping," which is available from Sawtooth Software.

Most perceptual mapping studies in marketing have three objectives. The first is to learn how products in a class are perceived with respect to their strengths, weaknesses, and similarities with one another. But knowing all that doesn't do you much good by itself. The second objective is to find out what it is that potential buyers want. Having those two parts, the third objective is to learn how to produce or modify a product to maximize its appeal to the target population of potential buyers.

These are "study objectives" expressed in marketing language. They're expressed a little differently in engineering or technical language. First, we construct a product space which is a geometric representation of customer's perceptions of products. Next, we need to estimate the density of demand throughout the space. Then, we need a way of predicting preferences for new or modified products. If we accomplish these objectives, we should be able to position products to perform most effectively in the marketplace.

Figure 1 is a map from my 1971 article in the Journal of Marketing Research, which shows the political landscape in 1968. If you can remember, in 1968 some prominent political

Figure 1

The Political Space, 1968

figures were Lemay, Reagan, Nixon, Eisenhower, Lindsay, Rockefeller, Percy, Kennedy, Humphrey, Johnson, and some non-candidates, including Martin Luther King and Stokley Carmichael. The horizontal dimension reflects the perceived Liberal/Conservative position of each political figure.

The vertical dimension is less clear, but seems to range from Anti-Administration to Pro-Administration. Stokley Carmichael, in the upper-left quadrant, was a dissenter, and George Wallace, in the upper-right quadrant was also dissenter but for different reasons. Johnson appears in the lower half of the map because he was the administration. Likewise, Humphrey, the Vice-President, and Kennedy, a former liberal President, also fall in the lower-left quadrant.

Political figures seen as similar lie close together and those seen as different lie far apart. Each "arrow" or vector represents an attribute. Consider the vector in the upper-left quadrant that says "withdraw Viet Nam." The public figures are positioned so that the farther out they are in the direction of that vector, the more they are seen to agree with that sentiment. Stokley Carmichael was thought to agree most strongly, and LeMay was seen to agree least.

Consider another vector: "income tax unfair." Wallace was seen as agreeing most strongly with that sentiment, and Johnson and Humphrey were seen as least in agreement.

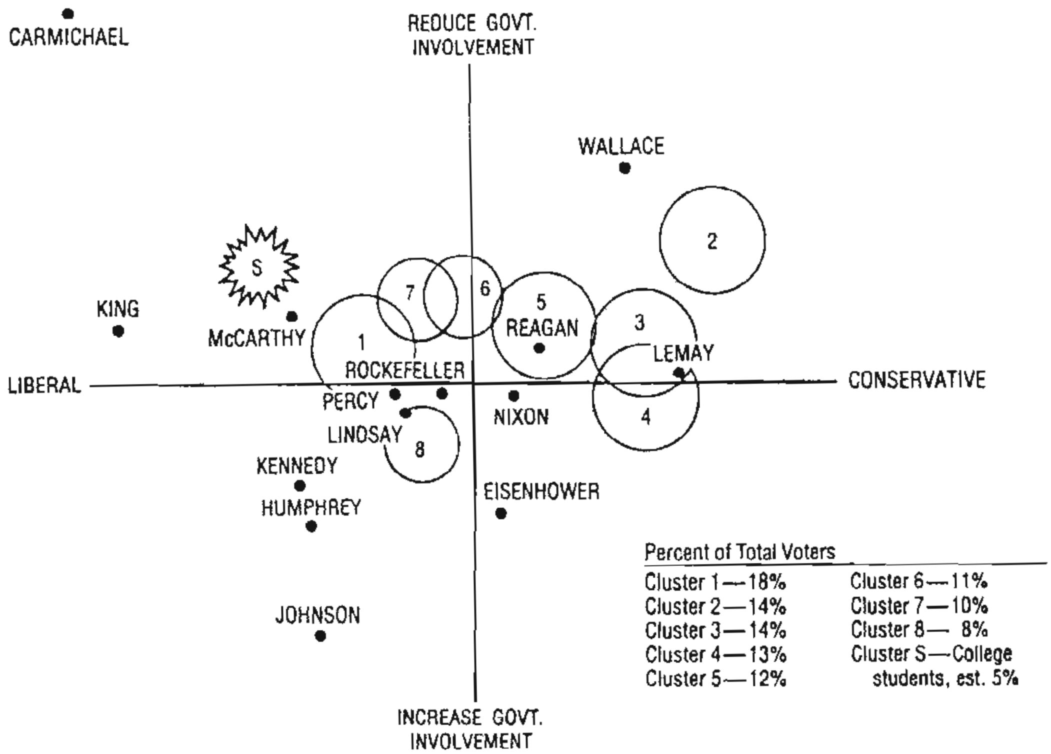
Every product has a location and every attribute has a direction in such a space. Attributes on which products differ strongly are represented by long vectors and those with less strong differences are represented by short vectors.

Hubert Humphrey, for example, might have used information like this to help reposition himself to appeal to more voters. This map could tell how he was seen on these issues, but it is not clear from this information alone how he might have had to change in order achieve that goal.

The ideal point clusters of voters are included in Figure 2. This is a very primitive representation, but it does convey an intuitive feeling for what voters desired.

Figure 2

Voter Segment Positions Relative to Political Figures



The average voter felt that his ideal candidate would be about in the center. That's quite a bit toward the political right and to the "north" of Humphrey. Had Humphrey wanted to be elected, he might have considered such attributes as:

"the income tax is unfair"
"farm subsidies should be eliminated"
"foreign aid has been squandered"
"everyone should be allowed to own his own gun"

If he could have found it in his heart to come out in favor of any of these issues, he would have been perceived as farther toward the northwest, and history might have been different.

There are three important methodological distinctions among ways of making maps. First, two kinds of data are frequently used to make maps: similarities data and attribute data. The APM System uses attribute data, even though similarities data can be more appropriate in some circumstances.

Similarities data are more appropriate for categories where the products are distinguished by nonverbal issues, such as tastes, odors, or aesthetics. In such cases it can be more appropriate to have respondents judge overall similarities among products, and attribute data have limited usefulness with such product categories. Another advantage of similarities data is not having to know before hand what the important product attributes are.

However, attribute data do have some significant advantages. Rating products on attributes is an activity that respondents seem able to perform easily and naturally. Also, researchers find such data inherently useful, and attribute data are already collected in many market research studies. Since the data are already available, it makes sense to use them for making maps.

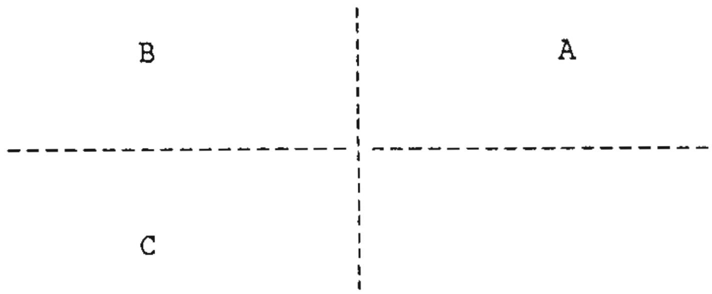
Given that attribute data are to be used, the second methodological distinction is whether to use factor analysis or discriminant analysis. Factor analysis has the advantage of producing dimensions that are purer and easier to interpret than discriminant analysis.

However, factor analysis usually produces more dimensions than discriminant analysis. We want to create an easy-to-look-at picture -- a two-dimensional representation on one piece of paper. Discriminant analysis will produce two dimensions containing more information than any other two-dimensional space. In a discriminant map there are typically several clusters of attribute vectors that would form separate dimensions if factor analysis were used.

The third methodological distinction is the way ideal points are handled; there are two ways to do this. It can be done implicitly by asking respondents which products they prefer, and then, based upon an existing map, inferring where their ideal points must be in order to support those preferences. Alternatively, it can be done explicitly by asking the respondent to identify his preferred (or "ideal") level of each attribute.

An implicit example appears in Figure 3, which shows locations of products A, B and C. If the respondent prefers product A to product B, we can infer that his ideal point lies in the right side of the space, which is the region closer to A than B. Furthermore, if he prefers B to C, we can constrain the region containing his ideal point to the upper-right quadrant, but that's still a broad area.

Figure 3



If we use just ordinal preference information, as in this example, the location of an ideal point will seldom be confined to a small region. Additionally, if there are many preference judgments, it is likely that they will imply contradictory locations, and some way must be chosen to resolve those contradictions.

Alternatively, we can offer a rating scale, such as "Liberal vs. Conservative" and ask people where they would place candidates A, B, and C on the scale, and to describe the level of their "ideal" candidate on the same scale. By treating the ideal as though it were just another product, it can be positioned in the map without special procedures. That is the procedure used by the APM System.

To summarize, the APM System uses attribute data, multiple discriminant analysis, and explicitly described ideal levels.

The APM System handles up to 30 products and up to 50 attributes. No respondent could provide good ratings for 30 products on 50 attributes. However, almost every researcher can think of 50 attributes for a category, and almost every client has at least 30 products that he wants to know about. We call our approach "Adaptive Perceptual Mapping" because the interview "adapts" to focus on subsets of products and attributes that make the most sense for that respondent.

The APM System has three parts: a questionnaire, a mapping capability, and a simulation capability. I'll describe each of these, starting with the questionnaire.

The questionnaire section lets the researcher compose a computer-administered interview, and in turn administers it to respondents. The interview has several well-defined sections.

The first section of the interview exposes the respondent to all of the products and asks a brief question about each. This question is typically concerned with familiarity. The responses are used, together with decision rules specified by the researcher, to select a subset of products for the respondent to consider in the remainder of the interview.

Next, each attribute is rated for importance, and decision rules specified by the researcher are used in conjunction with importance ratings to select a subset of attributes to be considered in the remainder of the interview.

In a product rating section, each selected product is rated on each selected attribute. The degree of selection is, of course, controlled by the decision rules supplied by the researcher.

Finally, there is a preference section in which products are presented two at a time and paired-comparison preference questions are asked.

A product familiarity question appears in Figure 4, where familiarity with Lee Iacocca is being rated on a five-point scale. Almost any scale or question could be used, and the researcher has control over where things appear on the screen. For instance, you could alternatively show all of the products on the screen at the same time and have the respondent select them in order of familiarity.

Figure 4

Suppose you were asked to give a detailed description of the background and political views of this person:

Lee Iacocca

Would you be able to do so? Please indicate your degree of familiarity with this person by typing a number from the scale below:

5 = Very familiar

4 = Somewhat familiar

3 = Only vaguely familiar

2 = I recognize this name, but that's all

1 = I don't recognize this name

A typical attribute importance question is shown in Figure 5. Almost any format can be used for these questions as well.

Figure 5

Suppose you were evaluating a candidate for an important national office, such as that of President.

How important would it be to you that a candidate be like this?

Conservative

5 = Extremely important to have this

4 = Somewhat important to have this

3 = This isn't really important

2 = Somewhat important NOT to have this

1 = Extremely important NOT to have this

After the respondent has considered all the products and all the attributes, making just one response per item, we collect ratings of a subset of products on a subset of attributes. Figure 6 shows selected products (Bush, Kennedy, Hart, Helms, and "Ideal") as they are being rated on analog scales.

Figure 6

How well does this characteristic describe the person below?

In favor of stronger defense

	Extremely Untrue	Extremely True
George Bush		
Jesse Jackson		
Ronald Reagan		
Lee Iacocca		
Your IDEAL candidate		

Use the arrow keys (<- and ->) to move the block to indicate your answer. Then press ENTER. Press "x" to back up.

We ask: "How well does this characteristic (`Conservative') describe this person?" At first, only Bush's name appears on the screen, and the respondent's task is to rate him. Then the next name appears underneath Bush's, and so on. Rather than an analog scale (I find it maddening to move the cursor across the screen slowly with the arrow keys), you can ask for a single-digit response. The objective is to get comparative ratings for the products on each attribute.

Next, preference questions are asked in which the products are compared to one another, two at a time. Again, the nature of the question is entirely up to the researcher. In Figure 7 we've asked: "if you could split your vote into 100 parts, how many parts would you give to the candidate on top?" You could also use a five-point scale, where 5 means "Strongly prefer the one on the top" and 1 means "Strongly prefer the one on the bottom."

Figure 7

Suppose these persons were running for the office of President.

How strongly would you prefer THIS person-> **Lee Iacocca**

compared to this person-------------------> **Jeane Kirkpatrick**

Suppose that you could split your vote into 100 parts, rather than having to give it all to one person. If you could do so, what percent of your vote would you give to the person at the TOP?

Type a number between 0 and 100 and then press ENTER.

Type "x" if you want to go back and change an earlier vote.

That's how the interview goes. The researcher has complete control of the screen. He can put questions anywhere, ask them in any way, and phrase them as he likes. The questionnaire can be in any language that uses Roman characters, a plus when doing international research. The researcher can choose the amount of customization of the interview. Respondents can be asked to rate all the products and all the attributes, or just some. Items can be randomized in many ways. The researcher can "force" specific products or attributes for each respondent, either unconditionally or conditionally (based on his own ratings for those items).

My last point about the APM System interview is one of the more important ones. Those of you familiar with the Ci2 System will recognize that all of these questions could also be written in Ci2. The first APM System study was, in fact, done using Ci2 since the APM questionnaire module wasn't ready yet. The Ci2 questionnaire required nearly 1000 questions and more than two days to "program." With the APM System, it took less than an hour to construct an equivalent interview. The difference is that the APM System has powerful "macro" instructions that let you specify vast portions of the questionnaire with very little effort.

The second part of the APM System does mapping. First, let me give a brief description of an important difference between discriminant analysis and factor analysis. For those of you who are not statisticians, there is a statistic called an "F ratio" which is a measurement of the size of the differences <u>between</u> things, compared to differences <u>within</u> things.

For example, suppose we had several respondents rate several political candidates on an attribute such as "conservative." We could compute the mean for each candidate to learn how different the candidates are seen to be from one another. The amount of difference between candidates might be expressed in terms of the variance of their means.

We could also see how much disagreement there is among different respondents' ratings of each candidate. For instance, if a candidate were rated on a ten-point scale with half of his ratings being 1 and half being 10, we would conclude that there's quite a lot of disagreement among raters. But, if he's rated 6,7,7,6,6,7,7,6 there isn't very much disagreement at all. The amount of disagreement among ratings of the same candidate might be expressed in terms of the variance of those ratings.

The F ratio is just the variance of the means for different products, divided by the variance among ratings of the same product. It's a measurement of the difference between products, where the unit of measurement is the amount of error variance that arises from different raters disagreeing when rating the same products.

A fundamental aspect of our approach to mapping is to concentrate on attributes on which products differ from one another, in ways that different raters agree upon. Attributes with high F ratios satisfy both of these criteria.

If you had ratings of political candidates on 50 attributes, you might compute the F ratio for each attribute, pick the attribute with the highest F ratio (the one on which the candidates are most markedly and universally perceived to differ), and make that attribute the first dimension. Then you might look for another attribute that also has a high F ratio but which is different in meaning from the first, and make that the second dimension. The resulting two-dimensional space might contain most of the information.

Discriminant analysis goes one step better. Instead of scanning all of the attributes to find the one with the highest F ratio, it finds that weighted combination of all the attributes that would produce the highest F ratio, and that combination becomes the first dimension. Then it finds that weighted combination of the attributes, uncorrelated with the first combination, which would produce the highest F ratio. That's the second dimension, etc.

The approach to discriminant analysis that we use is somewhat nontraditional, and follows my paper, "Multiple discriminant Analysis: Marketing Research Applications" published in Multivariate Methods for Market and Survey Research, edited by Jagdish Sheth and printed by the American Marketing Association in 1977.

The Mapping module lets you specify several different options. After you've made a map, you can display many kinds of results. In Figure 8, I've shown how our tiny group from southern Idaho viewed the world. On the horizontal dimension we have on the right-hand side, Reagan, Falwell, Kirkpatrick, Helms and Bush. On the left, we have Cuomo, Jackson, Kennedy, Hart, and Iacocca.

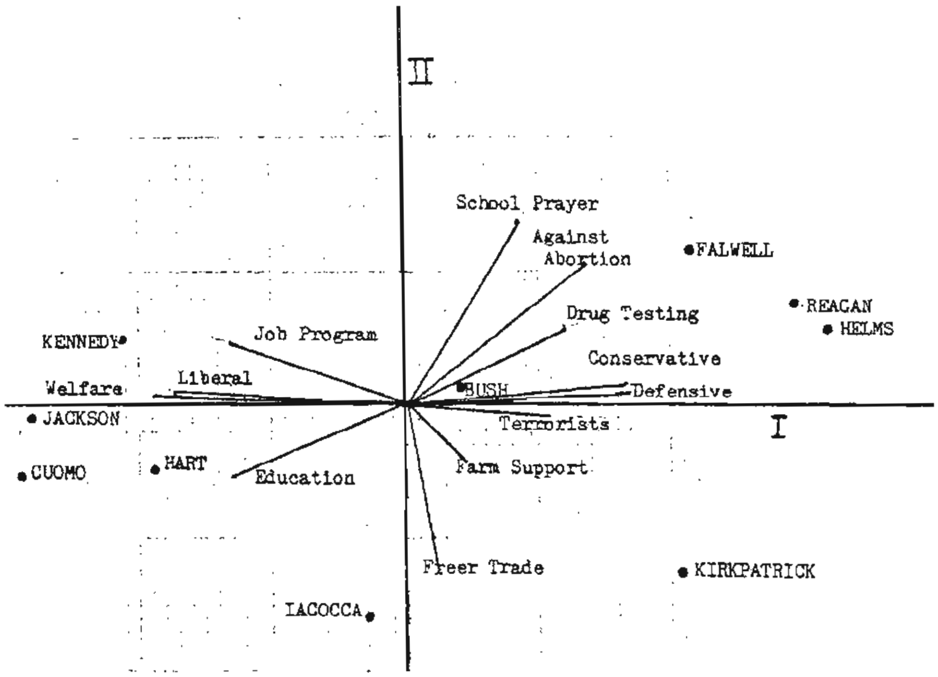
Figure 8

Candidates

	I	II
Reagan	1.47	.38
Bush	.17	.06
Falwell	1.09	.58
Kirkpatrick	1.20	-.66
Helms	1.61	.38
Cuomo	-1.45	-.27
Jackson	-1.42	-.05
Kennedy	-1.12	.23
Hart	- .95	-.25
Iacocca	-.15	-.80

Issues

	I	II
Liberal	-.99	.03
Conservative	.83	.09
Defense	.83	.08
Welfare	-.92	.02
Free trade	.12	-.60
Firmer/terrorists	.51	-.02
Education	-.65	-.26
Job programs	-.66	.23
Against abortion	.66	.53
School prayer	.42	.68
Drug testing	.58	.30
Farm Support	.22	-.21



This hand-drawn plot was made from data in the accompanying tables. The products and vectors are plotted using the numbers in the tables as coordinates.

To identify the dimensions, you can use correlations of each attribute with each dimension. For instance, the first dimension is correlated with "Liberal" (-.88) and "Conservative" (+0.83). According to Figure 8, conservatives are seen to be relatively more in favor of a stronger defense, drug testing, and school prayer and against abortion. Liberals are seen to be more in favor of job training, farm support programs, and increased welfare funding.

The location of each respondent's ideal point is also available. If you had 1000 respondents, you could add 1000 dots to this same map showing where each respondent's ideal candidate would lie. Figure 8 doesn't show ideal points because data from these eight respondents are too scanty. You can see how, given a piece of graph paper, a pencil and ruler, you could make a map. Of course, you can also use a plotter to make maps, but I find that doing it manually gives me more of a feeling for the product category.

The APM System can display any two-dimensional slice of the map on the computer screen, and rotate it any desired number of degrees. The products and the attributes move rigidly with respect to one another. Rotation is just a way of locating the horizontal and vertical axes wherever you would like.

In spite of their usefulness for many purposes, maps do present problems. First, maps are static. Maps are wonderful for seeing the way things are, but not very good for predicting the way things might be if you were to make a change. For example, if you were dealing with automobiles, and horsepower and acceleration were two attributes, they would probably be highly correlated and point in almost the same direction. Suppose we were to mount an advertising campaign to acquaint people with the differences in those attributes, and to explain why our product has fantastic acceleration without needing much horsepower. If we were successful we would have reduced the correlation between those attributes, and those vectors would no longer point the same way. An entirely new dimension may even be created. So, if you make a change, it's hard to predict what that change will do to the map.

Second, the map shows average perceptions, but people differ from one another in their perceptions. It's possible, of course, to map subsets of respondents. You can make a map based on the whole population, then show where products are as seen by rich people, poor people, men and women, and so on. That can help in reducing the loss of information that comes from looking just at averages.

But being restricted to averages does cause a problem, because we want to study individual ideal points; a map based on average perceptions is often incompatible with individual ideal points. There's no good way to put a respondent's ideal point into a map that's in conflict with his own perceptions. The use of maps that I've been describing, where you examine product and ideal points to see how a product might best be positioned, presumes that you can locate individual ideal points in that space. This may not be true.

A further point is that it's very hard to tell how much movement is enough. Let's suppose that we're Humphrey and that we want to move three inches to the northwest. How do we know we're moving just three inches, when perhaps two inches is not enough and four inches is too far? These problems lead to the notion that one might need something beyond maps to estimate what will happen in the market when you change a product's position.

The APM System offers such a market simulation capability. We use factor analysis to create a product space for each respondent separately, containing his product perceptions and ideal point. We assume that his liking for products differs inversely with its distance from his ideal point; he will prefer the product closest to his ideal, the next closest will be his second choice, etc. We also estimate a weight for each dimension so that his preference contours are as nearly circular as possible.

We convert his distances to shares of preference, and then aggregate shares of preference over respondents. That gives us preference shares for the "base case" in which products are as currently perceived.

Next, we conduct a series of simulations in which we reposition products in various ways. In each simulation we (1) change each respondent's perceptual data, (2) recompute product locations in his space, (3) remeasure the distances, (4) get new shares of preference, and (5) cumulate the results over respondents.

Every product we wish to simulate may not have been rated by every respondent, so there are likely to be some missing ratings. These can be handled in various ways. For instance, if a respondent didn't rate product X but we want X in the simulation, we can (1) ignore X for that respondent, (2) substitute average ratings for those who did rate X, (3) use specifications that we as the researcher enter separately, or (4) specify proxies for X. (If X is like Y, and the respondent rated Y but didn't rate X, then we can use his ratings of Y as proxies for those of X.)

The base case simulation result for this eight-respondent, Southern Idaho market is shown in Figure 9. We find that Gary Hart, for example, has a share of preference of 13.6.

Figure 9

SHARE OF PREFERENCE SIMULATION

	Person	Share of Preference
1	Ronald Reagan	16.0
2	George Bush	9.5
3	Jerry Falwell	7.4
4	Jeanne Kirkpatrick	12.3
5	Jesse Helms	3.9
6	Lee Iacocca	7.3
7	Jesse Jackson	10.3
8	Edward "Ted" Kennedy	12.4
9	Gary Hart	13.6

We can make changes in a product's perceived position simply by entering product number, attribute number, and the desired change. In Figure 10 we show the results of several simulations. In each simulation Hart was increased 10 points (on a 40 point scale) on a different attribute. The greatest gain in share of preference occurred with an increase in perceived "fiscal responsibility."

Figure 10

SIMULATING CHANGES IN HART'S POSITION

	(%)
Base case share of preference	13.6
Fiscally responsible (+10 points)	14.4
Balanced budget (+10 points)	14.2
Freer Trade (+10 points)	14.0
Firmer/terrorists (+10 points)	14.0
Lower taxes (+10 points)	14.0
Stronger defense (+10 points)	13.9

If this were a representative set of data, one might conclude that the most beneficial change for Hart would be an increase in perceived fiscal responsibility.

In summary, the APM System has three parts. The first part is a questionnaire module, featuring a computer-administered interview where the researcher has great latitude in phrasing questions, and in which the respondent may be asked only to rate familiar products on attributes important to him. An ideal level for each attribute can also be described explicitly. The second part is a mapping capability that uses discriminant analysis to locate products and ideal points. The third part is a simulation module, which models preference for each respondent and aggregates shares of preference over individuals to estimate shares of preference for new or modified products.

HOW TO SELL PERCEPTUAL MAPPING

by
Michael H. Baumgardner
Burke Marketing Research
Cincinnati, OH

In looking at the title of my talk, it occurred to me that we don't think in terms of selling perceptual mapping like it's a product we keep in our hip pocket and try to peddle to companies. Instead, it's a tool that any good marketing research company uses in selling solutions to marketing problems and answers to marketing questions. So the question is: When does perceptual mapping best represent the solution to a marketing problem? When you have a marketing problem for which perceptual mapping might be a useful tool, sell it as a solution.

We use perceptual mapping regularly at Burke, not so much for strategic purposes of doing research, but rather as a data summarization technique. In other words, in many studies you are collecting many attribute ratings on many different brands, and you are left with a great deal of information to look through in the form of a data table. It is very difficult to visualize that much information and pick up the essence of those numbers. We find that generating fairly simple perceptual maps is one way to take that information and put it in a picture form as a preliminary summary of the data.

On the other hand, there are marketing problems for which perceptual mapping might present a solution. As an example, I'll show a shortened presentation that we have done as a way of proposing perceptual mapping for a study.

When talking with a client, one of the best ways to communicate what we have to offer in the way of perceptual mapping is to develop an artificial case history. If the client manufactures packaged goods, we may have an example of frozen entrees; if the client manufactures computer equipment, we may have data processing equipment as our example. We tailor our example to the product category of interest. In this particular case history, the product category is frozen entrees; the client is committed to a $25 million advertising campaign; the key issue is whether the current positioning strategy is on target; and, if not, how can it be changed or enhanced. This client, then, has a positioning issue.

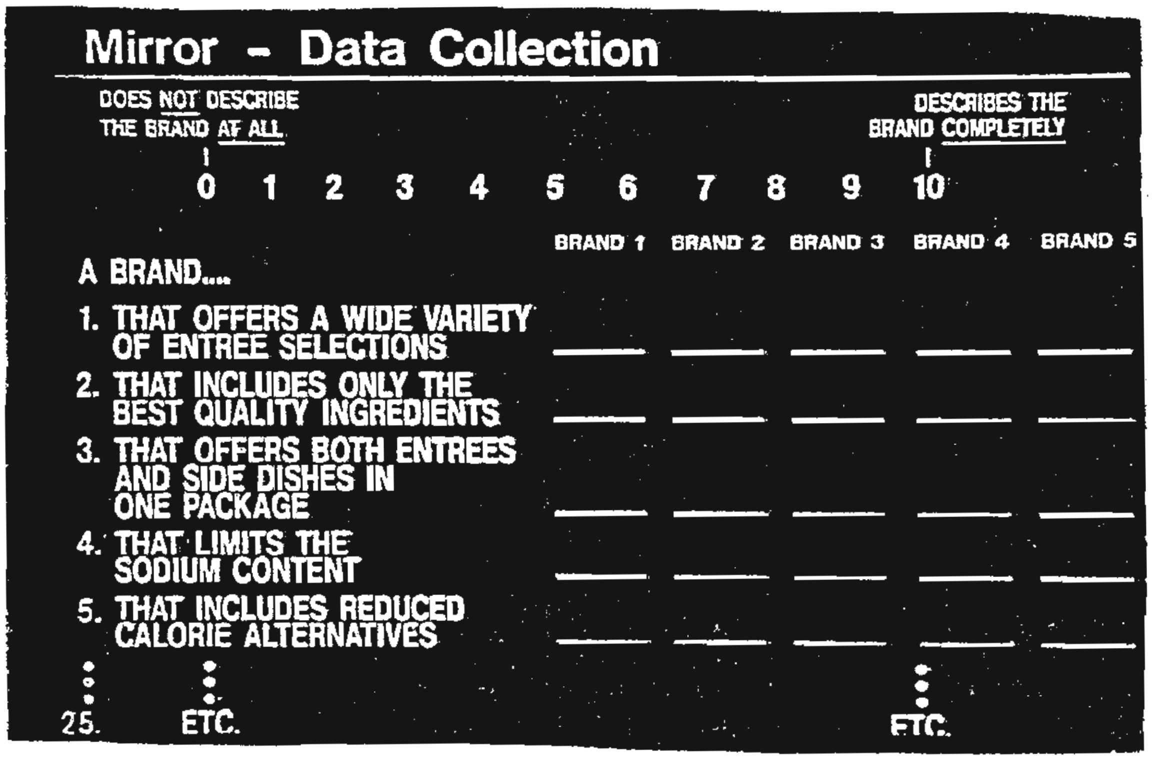
We walk through this particular case history to give the client an understanding of what he will gain from a study like this, and we point out some of the strategic directions that might be provided from the study. The objectives are to obtain, consolidate, and analyze information about the competition. In particular, we obtain attribute ratings of each competitor in the competitive framework. This is a particular approach to mapping that we use quite often at Burke, but there are a number of different types of mapping procedures that can be used.

We contrast the ratings for each competitor to the other competitors. In addition to getting information on the attributes and the brands, we always relate all information to some kind of critical behavior, for example, purchase intent.

We may look at these images of competitors across different buyer segments to derive some directions for the client in terms of repositioning their product, changing the strategy of their advertising campaign, or becoming efficient in improving their overall image.

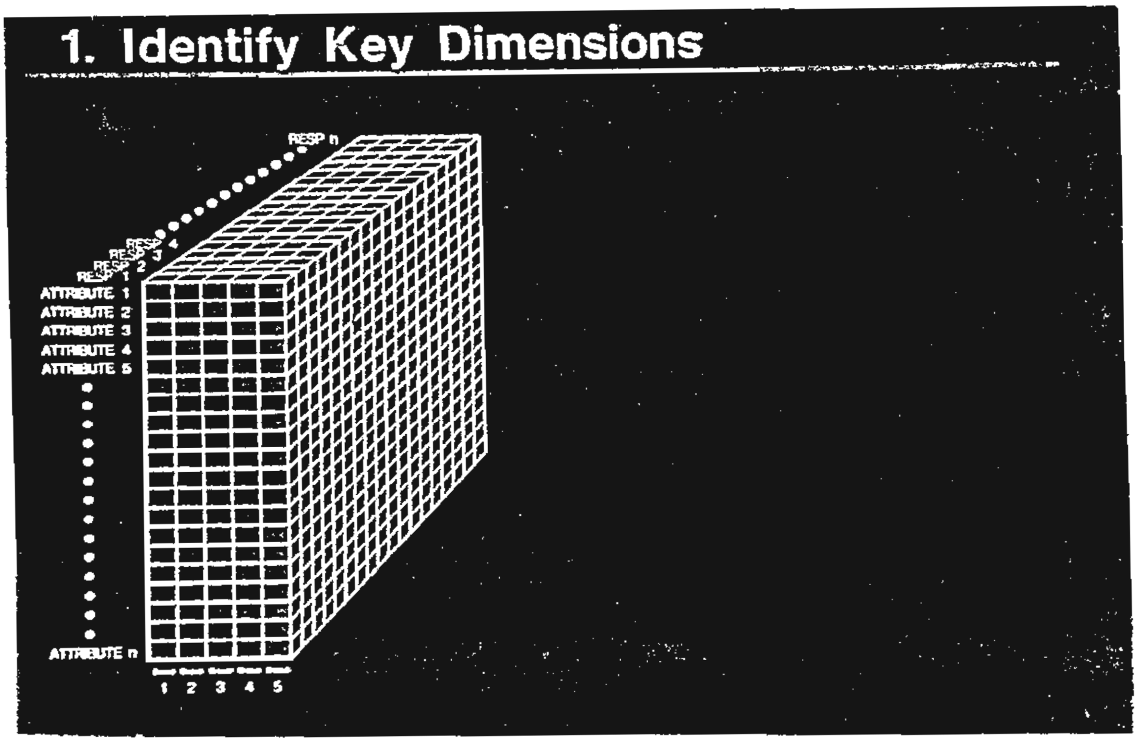
Figure 1 shows a particular approach called "Mirror" that we use at Burke for doing mapping. We start by taking all the brands of interest in the category, without going into detail about how many brands or attributes, and whether we use factor analysis, principal component analysis, or discriminant analysis. We gloss over much of this in a presentation to a client, because many of our clients are not statistically sophisticated enough to gain anything from that. Our goal is to quickly show the client the process and give him the end result to show the benefits to him.

Figure 1



We start with a number of brands in the category, take a number of attribute ratings on them, then arrive at data that can be visualized as a cube (Figure 2). From each respondent we have attribute ratings on a number of brands, which we reduce to the number of dimensions needed. If we have 50 attributes, we could have 50 dimensions. We want to reduce that to a smaller set through factor analysis.

Figure 2



In this particular example of frozen entrees, we found two important factors - "Breadth of Line" and "Premium Quality." Underneath each are the factors of some attributes that loaded highly on these two factors (Figure 3). After we have identified the various factors that underlie these attributes, we want to determine how important each factor is, relative to purchase interest. We perform regression analysis based on purchase-intent data for each brand. We regress factor scores on purchase intent.

Figure 3

I. Identify Key Dimensions - Continued

Factor 1

Breadth of Line

- OFFERS A WIDE VARIETY OF ENTREE SELECTIONS (.80)
- INCLUDES BEEF SELECTIONS (.69)
- INCLUDES POULTRY SELECTIONS (.65)
- HAS SELECTIONS THAT CHILDREN WILL EAT (.54)
- INCLUDES INTERNATIONAL SELECTIONS (.49)

Factor 2

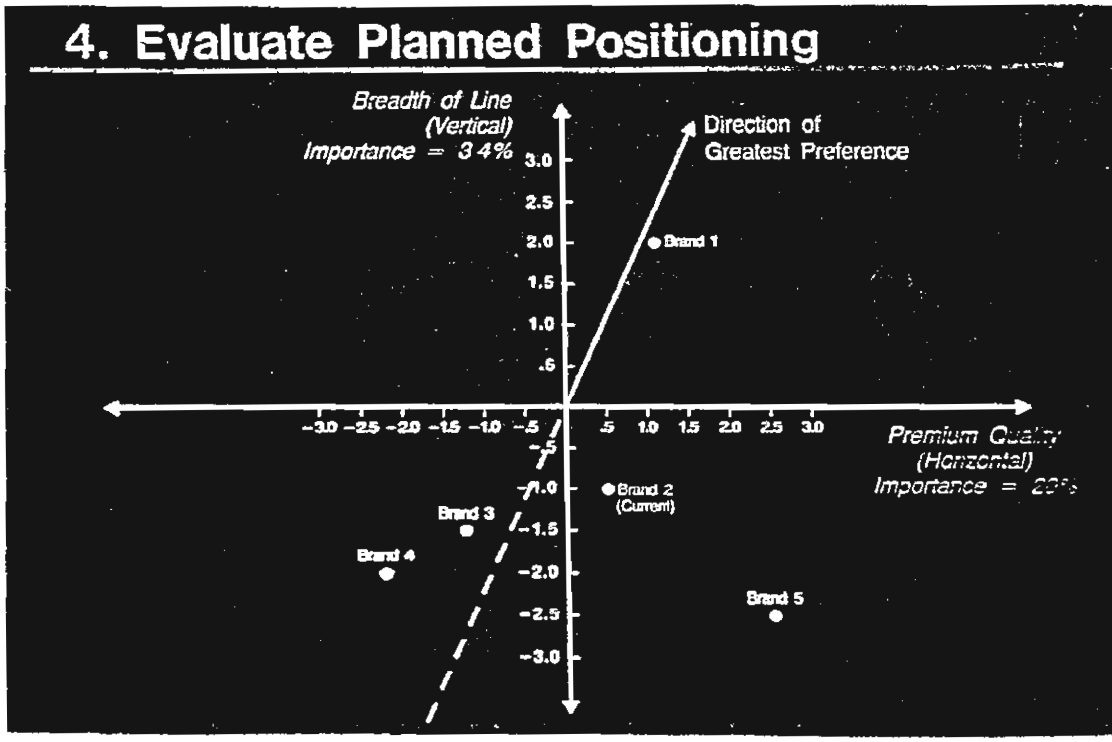
Premium Quality

- INCLUDES ONLY THE BEST QUALITY INGREDIENTS (.79)
- BASED ON CLASSIC RECIPES (.68)
- INCLUDES SHRIMP AND SCALLOP SELECTIONS (.62)
- IS A WELL-KNOWN, TRUSTED BRAND (.57)

This is a derived importance measure, as opposed to asking directly the importance of each attribute. In this particular example, the two most important factors are "Breadth of Line" and "Premium Quality." Although a number of other factors came out, we emphasize to a client prior to mapping that we can identify the important dimensions in his category as they relate to purchase interest. That itself is valuable information.

Let's look at the map in Figure 4. We have taken the two most important factors, Breadth of Line and Premium Quality, and plotted them based on factor scores. We plotted all brands based on correlations from the regression analysis as vectors in the direction of greatest preference. The vector for preference here is read the same as the vectors that Rich Johnson showed earlier. Notice that we have not plotted any attributes, but just labeled the dimensions "Breadth of Line" and "Premium Quality." We could overlay the dimensions onto this map, but for now we have reduced that set of attributes to an easier-to-handle size. The 10 attributes that fall under "Breadth of Line" most strongly characterize that dimension. We have reduced information through this map. For this particular client, we said that the general strategy of his advertising campaign was to emphasize premium quality ingredients.

Figure 4

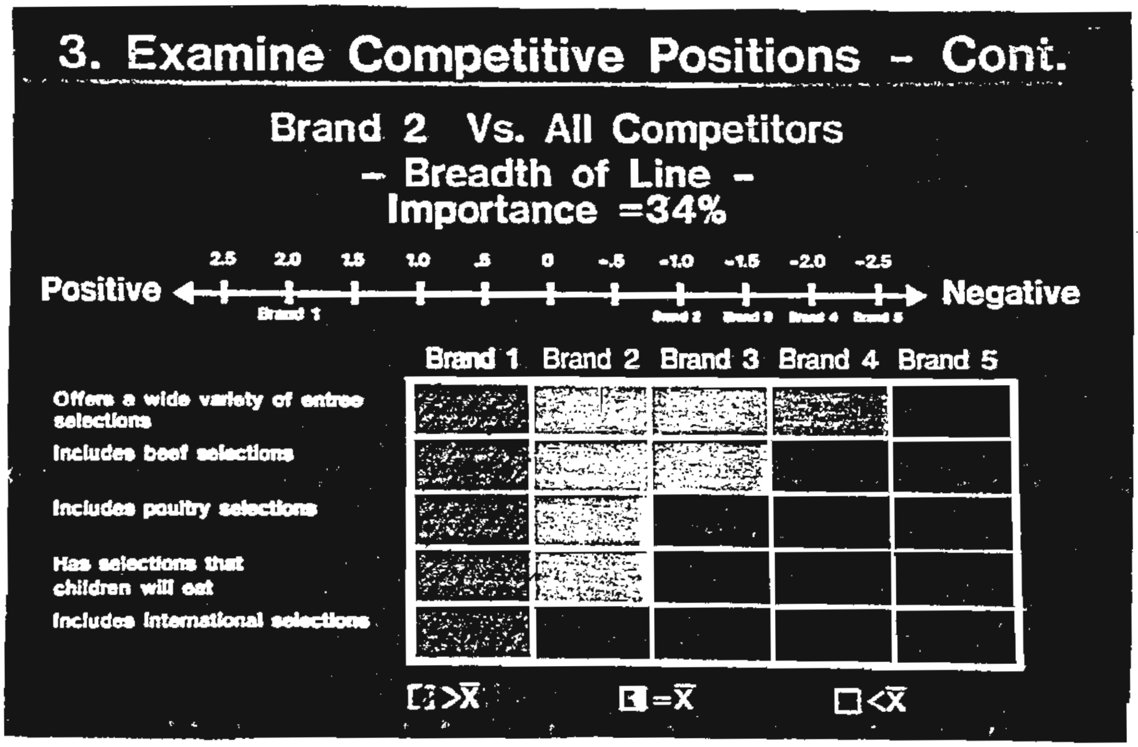


It is dangerous to move points around on a map, but we are going to do it here. In presenting perceptual mapping to a many clients, many like to push points around, and it's hard to stop that mentality. In this case, the implication is for "x" amount of movement on premium quality, which was the dimension they should emphasize in their advertising. You draw the perpendicular lines into the vector and look at the impact on preference for that much movement in quality. An alternative strategy would be to emphasize "Breadth of Line" and perhaps improve the breadth of the line. Assuming these are scaled equally, an equivalent movement on breadth of line should theoretically have a greater impact on preference.

In addition to this type of information, the key is that we have taken a relevant example to the client in a product category that has some meaning to him, and we have walked through the technique. We have glossed over the technical issues like factor analysis and regression analysis. At the end of the presentation we want to say, "Goodness, we can do all of this and save you making the kind of mistake you would have made here. You should consider this alternative strategy."

Of course, there are many other ways to look at these kinds of data. The chart in Figure 5, while not exactly a perceptual map, shows the client's brand versus all competitors on the breadth-of-line dimension, which had 34% of the importance that we allocated. You can see how the brands fall overall in terms of factor scores on the horizontal. Below are the attributes that load highly on this particular factor. You can also see how all the individual brands fall by color coding particular brands as higher than average, average, or below average. This is another way of extracting information from data.

Figure 5



Another issue is the sales presentation. Who should attend? Obviously, you want decision makers in the sales presentation. You always need to assume that the audience has never seen a perceptual map before. No matter how much they nod their heads while you explain how to draw a perpendicular to a vector, for instance, you have to assume that they don't really understand. Some people develop a mental block in dealing with points and vectors, and it's very hard to get across to them the idea of drawing perpendiculars as a way to read the maps. That requires educating clients who have not seen maps before. But once you do a few maps for a client and they learn how to read them, they like them and will come back for more. The first time, they draw perpendicular lines everywhere with a ruler. The second time, they start visualizing them, rather than having to draw perpendiculars. Before long, they understand how to look at and read maps, and that's when it becomes most useful for them.

Be sure to use a simple but relevant illustration while giving a presentation. Focus on the relationship of the map to the problem, rather than focusing on the technique itself. This relates back to my introductory idea that we don't sell perceptual mapping *per se*, but rather we sell perceptual mapping by showing how it fits into a problem, and how it can provide potential solutions to a problem.

Finally, you must understand the technical aspects in order to answer technical questions. Most audiences, however, will not be that interested in the technical details. If they are not interested in discriminant analysis versus factor analysis as a method, you should not try to impress them with that type of detail.

As for writing the proposal, give a brief technical description within the proposal and a detailed description of the technical details in the appendix. Generally, we include an example map in the written proposal that's tailored to the client's product category. We have a standard four-page example that we walk through that shows people how to read a perceptual map. We start with identifying points and vectors, then explain how to draw perpendicular lines and what they mean. We include that in our written proposals.

What are the most difficult aspects to communicate? There's no doubt that the most difficult aspects to communicate are the arrows, points, vectors, and brands. At first, it's difficult to grasp the concept.

Dangers of extrapolation occur once people get familiar with maps and comfortable reading them. Then you run into the problem of pushing points around without regard to the possibility that moving a point on the map may change the whole map or configuration in the whole space. It's something that you have to be very careful of, and there are many pitfalls in almost all approaches to it. Obviously, one of the purposes of doing a study like this is to give some direction and advice on which direction to go, so you can talk about general movements in particular directions without becoming specific.

People have trouble with the concept of the maps being relative, rather than absolute. For instance, the client's brand may be positioned well on a particular attribute like "good value for the money." The client will be very excited about that until you look at the mean ratings on a ten-point scale. Be careful that clients understand about relative positionings on the scales and not absolute positionings. The client may be the best on a particular attribute, but the whole set of brands may be performing poorly on that attribute.

The most common objection to using perceptual maps is that management won't understand them. To overcome that objection, you must educate those to whom you are selling the job. If they can grasp the perceptual mapping concept and see its benefits, it's much more likely that they will be willing to go to management and educate them on how to use it.

HOW TO SELL LARGE MARKET RESEARCH PROJECTS INTERNALLY WITHIN CLIENT COMPANIES

by
Edward (Ted) Evans
Ortho Consumer Products
Chevron Chemical Co.
San Francisco, CA

Mike Baumgardner has just provided us with an approach to selling perceptual mapping from a vendor's viewpoint. Let me share with you what happens on the other side of the fence; when we in the buying or client company try to sell these types of projects in-house. I will talk about selling large state-of-the-art market research projects of which perceptual mapping might be a part.

The projects that I will be referring to will cost the client company more than $100,000. Without getting into project details, this represents a project of such significant size that the divisional vice president cannot approve it at my company, but must seek the approval of our president.

The corporate environment in which I sell these kinds of projects is Chevron Corporation, a large international oil company that has among its holdings a wholly-owned subsidiary called Chevron Chemical Company (CCC). For the last five years, we have been undergoing some corporate cultural changes, referred to as manpower and expense cutbacks. To say the least, this type of climate is not always conducive to the support of large project expenditures that do not themselves assure high rates of return.

I work in the Ortho Consumer Products Division of CCC which generates somewhat less than 10% of CCC's revenue and less than 0.1% of Chevron Corporation's revenue. We are considered, however, by most in our company as the best marketer to the consumer because we enjoy a respectable market share in the lawn and garden chemical business, as well as a high consumer recognition of our brand "Ortho."

As part of selling a large consumer market research project in our company, it is necessary to assume that we do this only when the issues are serious and project goals and results must be understood and utilized at many levels of management. This is not a "nice-to-know-information" kind of project.

Project selling or communications needs to occur in three different areas in order to be successful:

- Within the business center, in my case the Ortho Division.

- Within the operating company that will ultimately approve the project because of the level of expenditure.

- Among vendors where communications will be vital during the approval process.

The characteristics of the kind of problem, opportunity, or project which can be given consideration for approval of large expenditures include: The division must be facing a "high-stakes" problem. By doing nothing, we may run a substantial risk. In addition, by taking the wrong approach the downside risk must be perceived as high, such as loss of the market share, revenue, and/or profits. As a result, we find ourselves in a situation where research-driven action offers a high reward opportunity, combined with all the elements that lead to controversy.

Much upfront, informal work is necessary before a project is officially proposed or approved by our management team. It is important to start this process well in advance of expected project approval or start-up. This is one of the most important parts of the internal selling process. Informal discussions or "pipe laying" will build a foundation of understanding that will be important to the proposal's reception by management. If reasonably done, this will flush out the objections that will need to be overcome before final approval. Dealing with these difficulties informally is far better than confrontations or emotion-packed management meetings.

As you are lobbying, do not forget to include all departments in your operating environment. If a large project is worth doing, it will be of interest to your associates in manufacturing, supply and distribution, product development and evaluation, finance, etc. Let them know what is occurring and what is at stake. Many market research projects have gone unapproved because of surprise objections, misunderstandings, or people feeling left out. By including them, you are likely to receive the team support your project needs.

The project steps as we normally think of them at Ortho start with defining project goals before establishing the budget. We can all remember times when it has happened the other way around. Too often, this results in compromised research with unsatisfactory results or budget overruns.

If you have done your upfront work well, the formal project proposal should be easier, but not necessarily a shoo-in. If there are to be difficulties, they will become apparent and you can find ways to overcome them. Including all department managers in the formal proposal has helped generate more rather than less support for the project.

We next solicit proposals from two to three vendors who are qualified. We do this objectively so that all who submit a proposal have a chance of being selected. Vendor selection occurs as a result of both their written and oral presentations.

After vendor selection, we seek formal approval from senior management (in this case, the president of Chevron Chemical). You may be sure that this is not the first time he has heard about this project. We normally select the supplier prior to final approval, because the vendor proposals play a significant role in project approval, as well as their selection. I doubt that all market research suppliers or consultants give this sufficient consideration.

Keep the size of the vendor sales presentations small to avoid the risk of distraction. On the other hand, it is important to include the key players facing the opportunity being pursued. We include both top and middle management and frequently someone from our advertising agency.

This is a very important session for the vendor, the client, and the project. All the vendors presenting will define (with our help) the goals for the project and the results the customer can or should expect. Methodology, however, is a different matter. While many suppliers know what should be done to accomplish the goals and provide the results, far fewer are capable of easily communicating them.

It is very important to avoid the jargon of your business. Use easily understood English (grade 10 or lower). While many have said this cannot be done, it is possible. For example, consider the jargon terms "Trade Off" or "Paired Comparisons." Why not talk about consumers or people making choices between things or ideas? Your audience need not understand all the statistical or quantitative ramifications of what will be done, but rather a simple outline of what will happen and some assurances that it will be done correctly. Offering to go into more detail with the client's technical people at another time will in most cases be satisfactory. If necessary, make two presentations; one for management and one for the technical people.

Talk to the customers (not a bad term considering we are in a selling situation) in terms that they can understand about <u>their</u> products and <u>their</u> kinds of consumers, not products and consumers that have been examined for other clients. My message is not a new one: Up-front work and preparation by vendors is the key to success.

When it comes to consumer market research, Ortho contracts on the outside for 100% of its requirements. The vendor characteristics we find important are:

- The ability of suppliers to communicate easily with a variety of client people with whom they must interact. Interpersonal skills and the ability to form a good working relationship with their people are critical. Having good quantitative technicians is usually not enough.

- Vendors who offer "state-of-the-art" expertise and alternatives, such as Sawtooth Software. We will not consider suppliers who are not up to speed.

- Good supplier reaction, follow-up, and feedback to customer questions. When lacking, it is painfully noticeable; when present, it means repeat business.

How the should the results of such a large, state-of-the-art project be used? By upper management? By middle management? For strategic purposes? As a tactical tool? My answer to all these questions is "YES."

HOW TO DESIGN A STUDY

by
Betty A. Sproule
Hewlett-Packard
Cupertino, CA

How to design a study - very carefully

I want to talk today about the critical step of designing a perceptual mapping study. I have tried to ferret out a few words to the wise from the experienced. My intent is to provide some general guidelines to help you get started, with emphasis on what to do and what to avoid.

First, you have to plan the entire study in advance. This doesn't mean the plan can't change as you work through the study, but you need to visualize the entire study all the way through. A person doesn't need a crystal ball to do this. The experienced researcher, and even the inexperienced, will be able to go through this phase. I often approach the plan through reverse engineering by starting at the back end. I ask: What are the decisions? How do they affect me? Will the data be given to the key decision makers as an oral presentation or written report? How should the data be formatted? What statistical analysis will be required? How will the study be fielded? How will respondents be identified? As for the questionnaire, how will the questions be phrased and in what order? In other words, you are considering in advance the entire design of the study.

Athletes in training are encouraged to visualize the entire process: the incredible amount of work that goes into training; the actual competitive event; hearing the crowd; feeling the competitors around them; running the race; seeing others fall behind; winning the race; and seeing themselves standing on the platform with the medal being placed around their necks while music plays in the background. In marketing research, you have to visualize the whole project and think of yourself as really being the force between the market place and the organization.

If marketing research can be likened to athletic competition, what would be the equivalent of standing on the winner's platform? Here is a real-life example that I want to share with you that happened at Hewlett-Packard soon after I arrived. I received an announcement to attend a marketing research presentation. The flyer had a footnote that read, "Attendance will be limited to the first 300 people who register." I should have known that this footnote was a tip-off to a really successful presentation. People arrived early; they stood in line. The had to sign in to attend, because attendance was limited. They vied for the top seats. I thought I was at a rock concert. They listened very intently, and there was a spontaneous standing ovation at the end of the presentation. Why? The study addressed issues of importance to those in attendance and did so in a clear and direct manner. The vice-president of marketing responsible for the project stood up and said, "It's been time and money well spent." He challenged each person in the room to consider how the results would impact what they were doing and what changes they would make; he suggested they confer with the marketing research personnel to make sure that they had not misinterpreted the results. Clearly this was an outstanding project that was well designed, well executed, and had the involvement and interest of many key people.

In designing any study, whether perceptual mapping or another methodology, you have to focus on achievable objectives. There are two key points: First, the achievable, and second, the objective. As for the achievable, perceptual mapping has the advantage that it's new and neat and everybody's into it. People feel perceptual mapping may be the ultimate marketing research tool that will solve all their problems. We have to be careful not to oversell perceptual mapping. People must have a realistic view of what the map will contribute to their marketing decisions.

Second, you have to ask: What are the objectives? I like the list of objectives that Rich Johnson put forward in his talk, because they are the traditional objectives that can be delivered by a perceptual mapping study. In identifying the key client groups within the organization, you can have many perceptual maps and many studies. But you have to ask: Who are the real individuals that we're talking about? Are we focusing more on image positioning or a product characteristic? Are we working with a broad spectrum of products, or ones narrowly defined? The schedule of activities, availability of results, and budget are other key items. Timing the availability of the data for decision making is an important issue about which people have to be very realistic.

You must achieve a consensus on objectives for research prior to initiation of a project. There are two schools of thought here. One says, "Let's really hammer out these objectives, and if we can't get agreement on the objectives, we can't get started." That's the purest way of approaching it. The other school of thought says, "Let's find a patron somewhere in the organization who has money, convince this person we need this study, get the ball rolling, and once the momentum starts, we'll get other people to come in on the project." This is the marketing research version of the debate: Should we get married and then live together, or should we live together and then get married? The key points are that you have to have some sense of objectives, know where you are in that process, and not be too naive.

Accurately defining the market defines the scope of your study. You could have a perceptual mapping study on almost any topic, but you must be certain that the focus of the study represents the issues needing to be covered. In defining the market, there are three things to consider. First, the competition. What is the range of products that are used and competitive to the product or industry being studied? When we do market research studies on Hewlett-Packard Plotters, one of our chief competitors is pencil and paper. If you look at a map among the set of plotters, you get one set of dimensions and attributes. If you look at plotters vs. printers, you get another. If you talk about pencil and paper and typewriters, you get a whole other subset of the competitive frame. Depending on the decisions to be made, you need to define the range of competitors. The extremes of the market very often give real insight into that definition. The market leaders usually can tell where the center of the map lies, but the extremes tell the boundaries and allow you to stretch the map.

A second point is to consider alternative applications of the product, because many times the market is defined by the product's use. An example would be Nestle's Quik. The primary use of Nestle's Quik is to stir it into milk. But there's a significant group of large volume users who use Nestle's Quik as an ingredient in a dry hot chocolate mix. The way they perceive the product and the competition for that product depends on the competitors and how the product is being used in each of the categories. Rich Johnson's beer study has a terrible flaw! It doesn't consider the use of beer as shampoo. Realistically, however, this flaw may not be major if this use is not part of the market.

You also need to designate a target audience. Candidly, we'll sell our products to anybody who will buy them. But when we do research, we want to ask: Who are those people who form our target audience? One of the advantages of a perceptual marketing program is that it allows people to speak about products with which they are familiar. You hope to identify some people who are familiar with at least some of the items in your study.

Developing the product factors or attributes is the key to the success of the study. In this case, quantity leads to quality. A study will be fatally flawed if you omit a dimension that is pivotal to describing or differentiating between products. The best way to avoid that problem is to generate many possible alternatives.

Creating the factors

Rich Johnson said that his APM program allowed for 50 factors, which I find a fairly conservative number. If you can't come up with at least 200, you haven't thought very hard. The question then becomes: How do you select from among items from this list? Developing product factors can be a fun and challenging task that should involve many people in your study. You should include all the users of your study in the creation of product factors. Ted Evans just talked about selling the study and involving people in the sales process. By involving people in the creation of factors, you have a follow-up activity. The people are involved once again by providing input into the study and contributing to its final output, which they have endorsed. Here is an opportunity to include the advertising agency as well, since there are not many chances to include them. Also include your public relations department. What your company actually says about products is quite important, especially in the written literature and documentation that's printed on the packages. Don't forget to tap your research and development and marketing departments.

In creating product factors, give considerable weight to the creative input from your marketing research personnel. Bill McLauchlan, Bruce and Richena Morrison, and I worked on a study at Brown and Williamson Tobacco where we had creative problem solving groups of marketing research people consider product attributes, as well as their own ideas and previous experience, to select factors for use in the study. The sessions were fun; I have very fond memories of them, and most proved very productive. When the final study was completed, a large number of factors had been generated by that creative process within the marketing research group.

Finally, read the background literature. The published literature helps in defining product factors, even from related categories, because many times some terms will be transportable across categories. Get especially articulate users who can talk about the category in creative ways and then review previous research within the organization. Perceptual mapping studies should be sold as an enhancement to previous research and not as a reinvention of the wheel.

Winnowing down the list

Having created a large list, you then have the problem of evaluating the product factors, trying to select those which will be the most useful in your final study. It's very important to come up with a manageable list. Too many factors make the results unwieldy and hard to interpret and manage.

I recommend a quantitative approach to winnowing the list. There are many ways to perform this evaluation function, but a rigorous quantitative approach is very important for the success of the study. Quite often time constraints make you feel you don't have time for a pretest, and you shortcut when you shouldn't. Initially, you need to establish what criteria will be used to select the factors. Since you've invited everyone to input their own prize factors, everyone has some vested interest here. You need to have objective criteria as to what factors to retain.

The ability to discriminate in a statistical sense is very important to provide coverage of the market, enabling you to adequately describe all the major players that you want in the study. Terms must be unambiguous. Just exactly what did you mean by "high-tech" or "conservative?" Terms have to have meaning or you have to do some development work to give them meaning. Terms also need to be actionable. They need to be usable by the marketing department or the R & D department to make changes in the product. Once you have developed your list of product factors and evaluated them down to a manageable list, you're ready to begin the field work, the implementation, and analysis of the study.

In the field, one important question to consider is: What can be done to insure respondent interest and involvement? You want your carefully identified respondents to be committed to the task. There are three key points to keep in mind: First, keep the task short. Even marketing researchers don't like to fill out long questionnaires. If you keep the task short, you're better able to get the involvement of many people. Second, gather usage information. If people use the product, they're committed to giving you information because they have a vested interest. Finally, make the respondent feel important. If people feel that their opinions are worthwhile they're more likely to participate. This feeling is more important than some of the actual incentives we give. Incentives should be used to create a perception by respondents that their time is important and, therefore, we are offering a token of our appreciation for their important time.

Still, there are differences of opinion on incentives. This was reinforced when I went to the Advertising Research Foundation conference the first week in March. A person in one of the booths claimed to have solved absolutely the problem of how to insure respondent interest in central location, mall interviews. Fantastic! He offered them a foot massage and said, "Come into my parlor; take off your shoes. Let me massage your feet with a little electronic vibrator while you fill in this short questionnaire." He claims it works all the time. On the other hand, most of us are not great on foot massage. There are other ways to insure respondent interest and involvement.

I've talked about visualizing the entire study, focusing on achievable objectives, and defining the market in the design of a study. How much is enough? The catchword is care. Successful marketing researchers do not merely look for good elements of design, they exercise such a level of care that they will not overlook any useful element.

HOW TO DESIGN A PERCEPTUAL MAPPING STUDY

by
Bill McLauchlan*
Burke Marketing Research
Cincinnati, OH

The important caveat associated with the design of mapping studies is that you don't necessarily design perceptual mapping studies. Instead, you design studies where perceptual mapping may be one of many tools that will be used in summarizing and presenting the data to the client organization. Underlying the study design process is a recognition that at some point you may want to look at a multidimensional representation of the data. It's important, therefore, to attend to a number of different features in designing a study that will provide the input for that multidimensional representation.

Some of the topics that I'll discuss will be redundant with those of other speakers, but I'll also cover the nuts and bolts of the practical limitations of what you can reasonably expect a respondent to do in an interview that includes brand attribute evaluations. Because brand ratings across attributes form the basis for mapping, the selections of the relevant brands and attributes are of critical importance in determining the ultimate success of the study. I use the term "brand" in a generic sense; we could be talking about consumer package goods or political figures.

In general, it's better to have too many than too few attributes. A missing important attribute may present a distorted picture of the marketplace. Because of the exclusion of some key dimension, the results may not bear a close resemblance to reality. The respondents should tell _us_ that we've been redundant in the generation of attributes, rather than our making _a priori_ decisions about what attributes seem similar. Such haphazard decisions can have disastrous consequences if carried to extremes.

*Now with McLauchlan & Associates, Cincinnati

In terms of brands, the same general guideline applies. It's better to have too many brands in the initial set than too few. Very often, in the area of consumer packaged goods, we need to remind our clients to talk about generics and store brands and the increasing role that they play in defining the marketplaces for brand name products. When defining the appropriate brands, it's also important to seriously consider what the relative competitive set is. Betty Sproule mentioned this in terms of Hewlett-Packard plotters vs. paper-and-pencil plotting. The definition of the relevant competitive set is a function of the study objectives.

If we're talking about presweetened breakfast cereal as a category, and we're working with a client who is developing a new entry, and he or she wants to understand how that new entry should be positioned in the marketplace, we may find ourselves in serious trouble if we obtain attribute ratings among just presweetened cereal brands. It may be that our new entry is likely to compete with other breakfast products: toaster waffles, hot cereals, as well as cold, ready-to-eat cereals, and other similar products. The competitive set may not be limited to the traditional category.

In addition, it might be important in defining our brand list to make flavor and variety distinctions. For instance, with soft drinks, the perceptions of diet 7-Up vs. regular 7-up may be very different, or diet Coke vs. classic Coke vs. New Coke. In the area of tobacco, Marlboro may have an image in a king-size box that differs from its image in a soft pack. It's important to keep those kinds of distinctions in mind when developing the brand list.

The single most important thing for suppliers to remember is that the client knows the category. The client is a true partner in designing the whole research program. Over time, we as supplier organizations can begin to develop a good understanding of the client's products, the categories in which they operate, the market leaders, and the market laggards. Initially, in getting involved with a new client, we need to rely on the client's expertise in helping us to develop that understanding of the category. In subsequent studies that partnership becomes even stronger.

How, then, do we develop these lists of products and attributes? One of the most important ways is the client-generated approach through such techniques as brain-storming, Synectics, and past research. The client has some understanding of the attributes that help to define the category in which the brand operates. As Betty Sproule indicated, ad agency input is relevant at this point as well. Beyond that, it may be necessary to go through a series of more "formal" types of attribute generation and evaluation processes.

The typical sequence in studies of this type starts with qualitative input that consists of focus groups, in-depth, one-on-one interviews that lead to attribute generation, as well as competitive set checks and market structure analyses. From these we learn how consumers define the category, and whether it's the same as our definition. The number and nature of the groups are a function of the specifics of the project, and it's difficult to generalize or be more specific than this. Between two and eight focus groups may be appropriate. Again, let consumers tell you about redundancies in the attribute generation process.

At Burke, we strongly recommend a type of pretest that differs from the kind usually associated with marketing research projects. It's a more quantitative review of 50 to 100 interviews to evaluate the attribute generation work. (More traditional pretests check questionnaire flow, time, and comprehension.) For instance, we may have our respondents go through a card sort of 200 to 300 attributes, sorting them into piles based on perceived similarities. We may label the categories "nutrition" or "taste" and ask them to sort the attributes into piles in which each seems to best fit. Or, it may be an unstructured sort where they put attributes into piles based on how similar they think they are.

In analyzing pretest data, recognizing that sample sizes are smaller than optimal in a statistical sense, you'll need to use some type of similarities analysis. It may be nothing more complex than a simple examination of the frequencies with which individual attributes are sorted with other attributes, or it may be more sophisticated, such as some sort of multidimensional scaling or principal components analysis to look at the attribute similarity. On that basis, we then have a more precise method or criterion for selecting attributes.

Ultimately, we want to go into the pretest with many, many attributes and come out with a much reduced list. We can take that list to the respondents in the main study with confidence that we understand the characteristics of the product category. Finally, we do a more traditional type of pretest in which we test a final version of the questionnaire on a small number of respondents to check the comprehensional aspects of the interview.

How do we typically structure a study that includes perceptual mapping as one of the expected outputs? We recommend that the brand attribute ratings be in the middle. We start with our traditional category awareness and usage information; brand awareness - aided and unaided; advertising awareness - aided and unaided; and brand usage information for the past three months or other relevant time-period.

Following the collection of category behavior, we start the actual brand attribute evaluation, where respondents rate the brands on the completed set of attributes.

Mike Baumgardner talked about how at Burke we also like to get a criterion measure, which we may relate back to the attribute evaluations in a regression sense. If it's appropriate, we'll get a purchase intent measure, overall opinion, willingness to consider, or some appropriate measure for the category.

We tend to discourage people from using psychographic batteries, unless they have been validated and proved reliable. We are often involved in studies where a client comes to us with a list of "lifestyle statements" that somebody wrote with no prior work to determine whether that battery is measuring what it's suppose to be measuring, or no prior evaluation of the reliability of the instrument. While we incorporate these batteries in studies from time to time, we recommend using highly validated and reliable instruments for this type of information. We want to preclude the possibility of doing cluster analyses on lifestyle statements that may result in segments with face validity, but from a psychometric point of view, may have no validity at all. At the end of the interview we collect demographic information and whatever else may be appropriate.

A key question for the brand attribute ratings is: Who's going to rate what? Rich Johnson described how in the APM System the algorithm, though fairly flexible, is one that gets at familiar brands. The question is: Should respondents rate familiar brands or all brands, provided the category is sufficiently small. How do we make that kind of decision? In general, at Burke, we recommend having respondents rate brands of which they are aware on an aided basis. We aid them with photographs, packages, brand lists, or rosters. The brands that they rate may include all that the respondent is aware of in some categories, or some subset that we control in other categories. For presweetened cereal, somebody might be aware of 75 different brands. We may pick five or eight from that massive roster for the respondent to actually rate.

Asking all brands may be appropriate if the client is considering moving into a category where he does not have a presence. The client may have strong brand name recognition in another category. He may not be in a new category, but may be looking to move into that category. A good example occurred in the last year when Coca Cola licensed their name to an apparel manufacturer to produce a line of Coca Cola clothing. Ordinarily, if we asked people what brands of jeans they were aware of, they'd answer Levis, Lee and other familiar brands. Most likely no one would answer "Coca Cola Jeans." It may be appropriate, in the context of the objectives of a study of that type, to force into the brand-rating set the name "Coca Cola Jeans" or the logo printed on the hip pocket.

At some point, whether we're dealing with aware-of brands, familiar brands, or all brands, we may need to reduce the number of brands that an individual is required to rate. If so, the data analysis will need to incorporate that subsetting process. Most of us are familiar with respondent scale bias: Some consumers are always "high end of the scale" users. Others are always "low end" users. Some answer "5" to every question, no matter what. A plan for adjusting the data for that respondent scale bias needs to be incorporated into the design. It must preserve the true differences that people see between brands, yet not distort the positioning of brands in a perceptual space.

A second type of bias can occur when you are working with a reduced set of brands. We call it an "evoked set bias." Brand A in the context of brands B and C might be very highly regarded. If that's your evoked set, and that's how you perceive the brands, that's how your ratings may fall out. Brand A in the context of brands D and E may be very poorly regarded. For someone rating brands in that evoked set, we would expect that to be reflected in their ratings. We need to plan to adjust the data, again, to preserve the differences people truly see between brands, while at the same time, removing the effects of the set of brands that the individual has worked with in arriving at those ratings. Basically, the rating that a brand receives is a function of the respondent's perception, not only of that brand, but the other brands that are being evaluated simultaneously.

A final point in the "who rates what" category relates to explicit and implicit ideal brand ratings. We typically collect an ideal brand rating in the attribute evaluation process.

How do you go about collecting the data? The options that we have available are the ones that we have available no matter what the type of study. These options include in-person mall intercepts, prerecruits, phone/mail, and phone/mail/phone studies. Thanks to Rich Johnson and his colleagues, we now have PCs available as well. The considerations that go into selecting a methodology are the same for mapping studies as for other studies. The considerations involve respondent qualifications, incidence, interview length, representativeness of the sample, complexity of the task for the respondent, timing, budgets, and the litany of other issues that surround the selection of a methodology for any type of study.

One of those considerations - interview length - poses the question: How long can we reasonably expect to keep these people around, given the various kinds of techniques? At Burke, we generally are willing to talk to people in the mall on an intercept basis for 20 minutes without an incentive, and up to 30 minutes with a two or three dollar incentive (Figure 1). In that kind environment, we expect our respondents to complete 12 self-administered ratings per minute using paper-and-pencil.

Figure 1

Issues in Deciding on Type of Data Collection

Interview Length

In-Person Mall Intercept:	Up to 20 minutes without incentive
	Up to 30 minutes with \$2 or \$3 incentive
	About 12 self-administered ratings per minute
In-Person Pre-Recruit:	Up to 60 minutes with \$25 or \$30 incentive
	About 12 self-administered per minute

Perhaps Rich Johnson can give some perspective on the ratings that surround the APM System and the time that we can reasonably expect people to spend doing those ratings. You'll have people who do 30 a minute, and people who do two a minute. But, on average, in planning what to incorporate in a study, we estimate 12.

So, in a 20-minute, non-incentive interview on the mall, we can expect a total of 240 ratings, if that's all we're doing in the interview. Obviously, that's not all we're doing. We'll be collecting other information as well. We need to put together an estimate of how long it takes to get the brand and category usage information and demographics. How many minutes remain, then, to get the ratings? That helps define the number of attributes and brands that we can expect people to cope with.

For in-person, prerecruit studies, we can keep people an hour with a 25 or 30 dollar incentive with, again, up to 12 self-administered ratings per minute. The overriding concern for this long an interview is the quality of data from someone doing 720 brand ratings.

In phone/mail interviews, we assume that people can do about 12 ratings per minute when we mail them an interview, although we have no control over this (Figure 2). While we could mail them an interview with 3000 brand attribute ratings, our termination rate will be excessively high if we proceed with that kind of interview.

Figure 2

Issues in Deciding on Type of Data Collection

Interview Length

Phone-Mail:	Assume about 12 ratings/minute Little control Length affects returns
Telephone:	About 6 ratings/minute Beyond 20 minutes will begin to see high term rates. Have the respondent rate all brands across an attribute; it's less boring than a brand across all attributes.

On the telephone, we expect about half as many ratings per minute as we get in an in-person, self-administered interview. Beyond 20 minutes, we begin to see very high termination rates. We also recommend in executing a brand attribute rating section over the telephone that respondents rate the brands across an attribute, rather than rating one brand across all attributes. It's considerably less boring and more involving for the respondent. At the same time, it forces some discrimination in the brand ratings that may not be apparent otherwise.

Advantages of the in-mall personal intercept include presenting the respondent with stimulus material, which may be a product, concept, brands, photographs, etc. We show and explain how to use the rating scale. For instance, we tell the respondent, "A seven means completely agree, and a five means this, and a three means this." Interviews are longer and the investigator can be available to clear up any confusion and minimize the item non-response. Our interviewers are instructed not to accept a "don't know" rating unless the respondent has said, "I don't know," ten times. The disadvantages include less representative samples, even in studies with large budgets and multiple locations. Other disadvantages are fewer rotations of brands and attributes. Because of the practical aspects of keeping track of interviews, we might have just four or five rotations in a study where the brands and attributes occur in different positions in the questionnaire. Another disadvantage is investigator bias. We have less field control and it tends to be more expensive.

For in-person, prerecruit studies, many of the advantages and disadvantages are the same as in a mall intercept study. A major disadvantage is its greater expense as a consequence of the high respondent fee. One advantage of a prerecruit study is that within limited areas, recruiting over the telephone may give a more representative sample than mall intercept studies.

Advantages of phone/mail studies are that you can see the rating scale, have a longer interview, and have better sample control. We can determine precisely to whom to mail the interview. We have no guarantee that the person we mail it to is the person who completes it, but when we put it in the mailbox, it's addressed to the right respondent. The disadvantage is that it's very difficult to present complex stimuli to respondents through the mail. Concept boards may be the extent of it. We are hampered by item non-response, and also mail interviews' inherent non-response bias.

Finally, for telephone studies, the advantages are that we can do unlimited rotations of brands and attributes. It's the least expensive way to collect such data. Field time tends to be shorter, and because it's done in-house, we have excellent field control as well. The disadvantage is that no scale representation is available by phone, unless you give the anchors on what might be a seven-point rating scale. You might constantly have to remind the respondent what those points represent.

Another important point is how to insure data integrity and quality in a study that involves mapping. These are the same kinds of things we like to incorporate in the design of any study: rotations of brands and attributes to insure that brands appear in different positions on a self-administered grid, and that attributes appear in a different order on that same self-administered grid. Keep the interview length as short as possible, not only within the constraints of what you can reasonably expect a person to do given a particular methodology of data collection, but also in terms of what you're asking them to do in the rest of that interview.

We typically validate between 10 and 15 percent on most studies, but 100 percent on legal studies. We want validation on all studies that are executed beyond our control, such as subcontracts to a field service. We get in touch with respondents to insure that they, in fact, participated.

For data entry verification, we insure that data coming from paper-and-pencil interviews are punched correctly, and that the ratings are the ones that the respondent actually gave.

We check for completes on self-administered ratings, and often don't allow "Don't knows." We try to force the respondent to give an answer by saying, "Well, you say you don't know, but based on your beliefs, your perceptions..." Then, if they're adamant about not knowing, fine.

Finally, for sample size considerations I'd like to say that a perceptual mapping study should have exactly 172 respondents, but I can't do that. It's a function of the number of brands, brand awareness, the number of subgroups of interest (which vary from study to study), the study budget, and the cost per interview. We have done these studies with sample sizes as small as 25 and larger than 1000. The study with 25 was in a category where the universe of potential respondents was 50. We got half and felt good about that. The studies with more than 1000 tend to be in large categories with large numbers of potential subgroups of interest.

ANALYZING AND INTERPRETING PERCEPTUAL MAPPING DATA

by
Paul N. Ries
Procter & Gamble
Cincinnati, OH

I would like to discuss briefly a number of analytical methods that have been used for perceptual mapping data, touching on some of the advantages and disadvantages of each. Time does not permit going into much detail on any one method, and I will try to keep the discussion as non-technical as possible. I will then conclude with some observations on methods for interpreting perceptual maps and making them more interpretable to nontechnical end users.

Analysis

The method used to analyze perceptual mapping data is dependent upon two major factors: The type of data collected, and the information we wish to represent in the map.

As mentioned previously, perceptual maps can be derived from three kinds of data:

- -Perceived similarities between products
- -Rank orders of preference by individual respondents
- -Ratings of association between brands and attributes

Similarities Data

Perceived similarity data are most useful when trying to map a product category which is new and unfamiliar, or when attributes are difficult to describe in words. For example, we might wish to map the category of patterned toilet tissues or a set of potential designs on bath towels, which might be used as promotional premiums. While the patterns and designs can be described in terms of how well they go with bathroom colors, etc., it is difficult to find words which differentiate them from a cognitive viewpoint. However, respondents can judge similarities of patterns and designs relatively easily.

There are a number of ways to collect similarity data. The ones with which I am most familiar are unconditional judgments of pairwise similarity (i.e., ratings of similarity on a categorical scale from "identical" to "totally different") or conditional rankings where one product is chosen as the "anchor" and the respondent ranks the other products in descending order of similarity to the "anchor."

Similarities judgments are mapped by means of a multi-dimensional scaling algorithm. The objective is to represent perceived dissimilarities between products as distances between product points in a space of minimum dimensionality. The relationship between the perceived dissimilarities and the distances may be one of having the same rank orders, in non-metric multi-dimensional scaling, or of being linearly related in the more traditional metric scaling procedures. Other algorithms allow non-linear but functionally defined relationships between the dissimilarities and distances. It is even possible to position the product points in non-Euclidean spaces.

For most practical marketing and product development mapping problems, non-metric scaling algorithms like KYST, from Bell Labs, or ALSCAL, now available in SAS, provide quite reasonable flexibility and are fairly easy to use. We have also worked some with a program called PROSCALE by Dave McKay at Indiana University and Gerry Zinnes at the University of Indiana. It provides for the possibility of differential certainty about the dissimilarities between various pairs of products, and can provide confidence intervals on the locations of the product points.

Regardless of the method used for multi-dimensional scaling, the output is a map of the products on unidentified dimensions. We simply know that the space has two, three, or more dimensions. For example, if we are mapping automobiles and find one direction in the space along which the products are in the order:

1. Rolls Royce
2. Mercedes
3. BMW
4. Lincoln
5. Oldsmobile
6. Yugo

we might infer that this direction is highly related to price and/or luxuriousness.

Even if we are not sure we know all of the attributes that might differentiate products in a particular category, we can list those which we think might be relevant and have respondents rate the products for association with these attributes. The resulting scale ratings can be located as vectors in the multi-dimensional scaling configuration using a program like the Bell Labs PREFMAP Phase IV or a multiple regression program, using the coordinates of the products as predictor variables and the attribute scale as the response variable.

Such external fitting of attribute vectors can be a considerable help in interpreting multi-dimensional scaling solutions. It is also possible to determine whether the products differ along some dimension not well related to any of our list of proposed attributes.

Preference Data

Although less frequently used in market research, maps can be produced directly from scale ratings or rank orders of preference by individual respondents or groups of respondents. I am not sure these should be called perceptual maps, but they do serve the same purpose and often look the same as maps produced from other types of data. There is a growing interest in such maps because they can be produced from scanner or diary-panel data that reveal a history of actual household purchase behavior. I have no direct experience with such applications, having used the technique only with responses to preference questions.

Preference maps can be prepared with the multi-dimensional scaling algorithms I mentioned earlier. Respondents are rows and products are columns of a "lower corner" matrix that is treated by row-conditional non-metric regressions. There are also newer algorithms, such as DeSarbo and Rao's GENFOLD2 with which I have no experience.

The result of multi-dimensional unfolding is a map in which both respondents and products are represented by points. The points are positioned so that the distances from each respondent point to the product points best reproduce that respondent's product preferences.

With such maps, there are no attributes involved. The problems of interpretation are the same as with multi-dimensional scaling of similarities data. My experience with earlier multi-dimensional unfolding algorithms is that the output configurations are highly dependent upon the sample used. Problems with degenerate solutions occur much more often than when using the same algorithms for multi-dimensional scaling.

Attribute Rating Data

Most perceptual mapping for marketing research is done with some form of ratings of the association between brands and attributes. Most commonly, respondents are asked to use some kind of categorical scale to assign a numerical grade for how well each attribute describes each brand. If each respondent is asked to rate a moderately large number of brands on a moderately large number of attributes, the resulting task can be formidable and respondent fatigue may affect both the completion rate of the questionnaire and the reliability of the data.

Before the interactive capability of the Sawtooth Software APM System to choose an appropriate subset of brands and attributes for each respondent, the most common solutions have been to limit severely the number of attributes and brands included in the study, or to break the questionnaire into arbitrary subsets and use different samples for each subset.

A different solution, which I have often used, is to collect only dichotomous (yes/no) data with a brand-by-attributes checklist format. This method permits fairly large numbers of brands and attributes without presenting too great a task to the respondent.

With scale rating data, the most frequently used mapping techniques have been principal component and multiple discriminant analysis. Multiple discriminant analysis is, of course, the procedure used in the APM System and has some very desirable properties. However, most of my experience and that of agencies and groups I have worked with has been with principal component analysis. Some of my colleagues have also used classical psychological factor analysis which differs from principal component analysis in the underlying model and, to some extent, in the computational procedure. The differences are beyond the scope of this discussion.

Principal component analysis is available in nearly all popular statistical analysis packages for both mainframe and microcomputers. A typical and versatile example is PROC Factor in mainframe and PC SAS, which also offers a wide variety of classical psychological factor analytic methods together with a variety of rotation options, yielding both orthogonal and oblique rotations.

In using principal component analysis, I have generally performed an initial transformation on the data like that used in the APM program, i.e., each respondent's ratings of the brands on each attribute are adjusted to a mean of zero. This is done to remove between-respondent variance and concentrate the analysis on the between-brands, within-respondent variance. The input data set is then a matrix with attributes as variables and as many observations as there are subjects multiplied by brands.

I also specify a covariance rather than correlation matrix as the first step in the principal component analysis. This option has the effect of focusing the analysis more on those attributes where respondents see larger differences in the brands. The effect is usually a solution of lower dimensionality than would be obtained by use of a correlation matrix, an effect similar to that achieved by multiple discriminant analysis.

With dichotomous check box kind of data I mentioned earlier, I have generally used either Correspondence Analysis or a variation sometimes called bi-plot analysis. Correspondence analysis was developed largely by Benzecri at the University of Paris in the 1960's, and has been used in European, particularly French, marketing research circles for the past two decades. However, Benzecri's work was not translated into English and the method was virtually unknown in the U.S. until the appearance of books in English by Greenacre and Nishisato.

Until recently, we used a program written in the early 1970's by a European employee who had studied under Benzecri. The procedure is now easily programmable using SAS PROC IML, on either mainframe of PC SAS. Time does not permit a detailed description of the technical details. Briefly, the matrix of frequencies of associating brand with attributes is treated as a type of contingency table, and the chi-square associated with the table is decomposed into components, just as in principal component analysis.

The result is a relatively low dimensional map; typically two or three dimensions will account for 70% or more of the total variance. For all the differences in data collection and analytical methodology, correspondence analysis of dichotomous rating data tends to produce aggregate maps very similar to those coming from principal component analysis of categorical scale rating data. The major advantage to this approach is the ability of respondents to rate more brands on more attributes than they possibly could when using categorical rating scales. The major disadvantage is that we cannot produce individual respondent level factor scores, only an aggregate solution.

Other algorithms and procedures for creating perceptual maps exist, many of them with particular advantages for very specific kinds of problems. I have mentioned briefly only those which seem to be most widely used in commercial market research applications. As a group, these methods seem to be quite robust, producing very similar results, at least in the first few dimensions extracted.

Number of Dimensions to Extract

I believe very strongly that perceptual maps should always be limited to a relatively small number of dimensions, probably no more than four. Even with three dimensions, we begin to lose the ease of visual interpretability which is so obvious in all of the two-dimensional maps typically used as illustrations in textbooks and popular articles.

Rotation

The APM System provides a very convenient facility for manual rotation of the axes of a perceptual map. Rotation is usually necessary to improve the interpretability of the map, even though most of the relevant information can be obtained by looking at the locations of the individual attribute vectors in an unrotated space. In practice, I have found that most marketing people have difficulty dealing with a large number of attribute vectors, particularly in spaces of more than two dimensions. They find it much easier to deal with the two, three, or four axes of the space and would like to have summary names attached to each.

With a two-dimensional map, manual rotation is usually quite easy. With three, it is possible, but may be time consuming and also somewhat confusing. I speak from experience with a program of my own that does much the same kind of rotation as the APM System, although at a considerably slower speed.

I find it advantageous to submit maps of three or more dimensions to an initial varimax rotation, before attempting improvements by manual rotation. Often the varimax rotation alone is adequate, although it should not be accepted without close examination. Most principal component analysis programs permit automatic rotation of the solutions using varimax or other rotation options. For solutions produced by Correspondence Analysis and other techniques, I have used one of several stand-alone rotation procedures, the easiest being one in the PC Statgraphics package. I believe that it might be very useful if a varimax rotation procedure could be built into the APM System.

Representing Respondent Preferences

Perceptual mapping, as I have described it up to this point, is a process of modeling consumer perceptions of the characteristics of brands in a product category. The one exception is the process of multi-dimensional unfolding which derives both brand positions and individual respondent ideal points from preference data. With all other techniques, the next modeling step proposes a causative relationship between these perceptions and consumers' attitudes toward or preferences for the brands.

A number of different models are used for this purpose, only a few of which I can mention in the time available. One way they may be classified is by whether they use respondents' direct ratings of an ideal brand on the attributes, or derive an ideal point from the locations of the existing brands and preference for, or overall opinion toward those brands. The APM System uses both ideal brand ratings and preference ratings.

They may also be classified by whether they assume: 1) the same perceptual structure <u>and</u> brand locations for all respondents; 2) the same perceptual structure but <u>different</u> brand locations for each respondent; and 3) assume <u>both</u> different structure and different brand locations for each respondent. The APM System falls in the latter category.

I tend to prefer derived ideal points because I have observed that most respondents, when asked to rate an ideal brand, will assign it to the top scale position for all "good" attributes and to the bottom scale position for all "bad" attributes. Only in rare cases will respondents choose some intermediate level of an attribute as "ideal." Yet in the type of minimum dimensional maps with which I work, most dimensions will have "good" attributes at both ends. Conversely, most high share brands tend to map into positions fairly distant from the origin, indicating distinctive strengths on some attributes, but not on others. Brands near the origin tend to be those with no distinctive characteristics.

The problem arises when we try to plot the ideal brand into a space derived for existing brands and it does not fit the structure. It wants to be at both ends of the dimensions, simultaneously, and usually ends up near the origin. It does not help in the aggregate map, to include the ideal brand in determination of the structure, because it will have relatively little influence, the structure being mainly determined by the much larger number of existing brands.

Most of my experience has been with models that:

- -Assume constant structure for all respondents
- -Assume either constant brand locations or allow for individual differences in location
- -Derive an ideal point location from preference data rather than using respondents ratings of an ideal brand

The first of these models was originally developed by Carroll and Chang at Bell Labs, and is incorporated into a program called PREFMAP.

PREFMAP includes four phases, three of which are ideal point models and the fourth a vector model, i.e., preference is represented as a vector from the origin pointing in a direction of maximally increasing preference. A separate point or vector can be derived for each respondent, but individual level data is often too unstable for the purpose, since an individual rarely has experience with more than a few of the brands. My own approach has been to cluster respondents on the basis of their patterns of preference for the brands, and then derive vectors or ideal points for the cluster centroids.

Vector models can also be fitted by multiple regression programs, using individual respondent level product positions as the predictor variables. Again, a vector can be derived for each respondent, but vectors representing cluster centroids tend to be more stable.

The APM System differs from all other ideal point models in that it uses individually derived structures for each respondent, with the "ideal" brand included together with a small number of most familiar real brands. Thus, the ideal brand has more of an opportunity to affect these individual structures. Further, the logistic regression procedure used to derive dimensional weights provides a measure (in the correlation coefficients) of how well the model does fit. I hope to use this model on some real data in the near future to become more familiar with it.

Simulation

The purpose of relating brand positions to preference is to permit estimation of what changes in preference might be affected by repositioning an existing brand or introducing a new brand into the category. Perceptual maps can provide some help in answering such questions, but there are potential problems.

In the conventional preference mapping procedures ideal points or preference vectors can be plotted into the aggregate map. For any one such point or vector, we can see on the map the direction of change that maximizes the increase in preference. We can evaluate a potential new brand positioning in terms of how close it is to the ideal point or how far out on the preference vector it projects, relative to other brands in the category. Computing the actual predicted change in preference requires plugging the positioning into some form of regression equation, and multiple ideal points or vectors must be accounted for by adding or weighing together many such predictions. Thus, the need for a "simulator."

The map, however, suggests what directional changes should be tried in the simulator. It confirms, in retrospect, the simulated changes in preference due to repositioning.

In the APM System, the ideal points used by the simulator do not exist in the space shown by the aggregate map. They exist in many different individual respondent spaces. The aggregate perceptual map provides relatively little guidance in how to reposition old brands or position new ones.

I have had no experience using the APM System simulator on real data, so I do not know how difficult it may be to find the directions to higher preference areas of the map. However, it may be more appropriate to use some kind of automated grid-search procedure than would be necessary with more conventional preference maps.

Segmentation

There are two forms of segmentation which may be used in conjunction with perceptual mapping: brand segmentation and consumer segmentation.

In brand segmentation, we are looking for sub-categories of brands, often designed for different usage purposes or with different in-use benefits. For example, it is well known that there are separate sub-categories of bar soaps, one of which offers deodorant protection, another cosmetic benefits, and even a newer sub-category of liquid soaps for hand washing and bath and shower usage.

These sub-categories will be characterized by clusters of brands, perceived as relatively similar to one another and different from brands in other sub-categories. They are usually relatively easy to detect using hierarchical clustering procedures based on the Euclidean distances between brands in the mapping space.

Consumer segmentation is a more difficult problem, and a subject unto itself. Many different variables can be used to cluster consumers into segments: demographics, personality and lifestyle variables, product usage habits, etc. In conjunction with mapping, I have often tried to segment respondents on the basis of differences in their patterns of attribute importance ratings and/or preferences for the brands. I have found that true clusters based on these variables rarely exist. Usually the clustering algorithm produces partitions of the total respondents into groups that differ in their preferences, but are highly unstable and data dependent.

These clusters are useful if you wish to combine respondents to gain stability in preference mapping operations. They maintain much of the individual level diversity of preferences without including as much random error. I would be very hesitant to represent them as naturally occurring sub-groups of individuals in the sense that we usually mean when talking about market segmentation.

Summary

I have outlined a few of the many options available for creating and using perceptual maps. Unavoidably, I have probably failed to mention some very useful techniques that many of you have used with success. In particular, I think of LINMAP developed in part by Allan Shocker, INDSCAL, and the closely related three-mode factor analysis of Tucker. Each has its place for specific individual problems. I have outlined a few techniques which have proven with time to be broadly applicable over a wide range of marketing problems. I have tried to position these techniques relative to those used in the APM System, which will become a valuable addition to our methodological arsenal.

ANALYSIS AND INTERPRETATION OF RESULTS

by
Paul Hase
Hase/Schannen Research
Princeton, NJ

When I accepted Rich Johnson's invitation to speak at today's conference, my initial reaction was to present a wide variety of maps, generated under varying technical conditions and for varying research objectives, and to discuss their interpretation and applications. However, to do so would not have provided any general overriding message about the topic.

My first hurdle was finding a title that would convey the fine line between my historical apprehension toward mapping and my generally optimistic view for the future. Some of the titles I considered included:

-Vectors, Vectors Everywhere, But In What Direction Do We Go?

-The Elusive Market Niche - Find Me a Way to Point the Way

-The Fifth Dimension: Singing Group or Subtle Insight?

-We Must Do More Than Map!

Obviously, my predisposition toward mapping has not been entirely positive.

One issue of concern is that mapping is often considered an objective of a study, as having a life or reason of its own. This has serious ramifications for us in marketing research. Just as I react strongly when someone says that he wants to do a "segmentation study," so, too, do I react strongly when someone says he wants to do a "mapping study." We do NOT do segmentation or mapping studies; we do market investigations, opportunity studies, or new product development/brand modification studies - studies with titles that connote research and business objectives. Segmentation analysis, perceptual maps, and other tools are lines of analysis, not types of studies. Using study titles such as "mapping study" overemphasizes techniques, rather than conveying that as marketing researchers we do research to solve business problems.

In their book, Applied Multidimensional Scaling (1972), Green and Rao apply 16 different multidimensional scaling algorithms (from a total of 33) to a given set of data. Their objective was to demonstrate various methodological differences in the scaling (mapping) approaches they evaluated.

In their introductory chapters they list three alternative approaches to the scaling of similarities data:

- Whether the respondent judges the objects, e.g., brands, on whatever criteria he wishes, or evaluates them on a prespecified set of items, e.g., rating scales

- Whether the data are scaled at the individual respondent level, or aggregated in some way

- Whether a metric or nonmetric mathematical algorithm is used to scale the data

For the scaling of preference data, they also list a similar set of issues that differentiate various approaches to the scaling/mapping problem. The scope of their exhaustive methodological examination is beyond the intent of this presentation.

Instead, I will delineate pragmatic issues and show how maps are built. This approach may seem simplistic, yet it illustrates what maps do and do not do for us. I also hope to dispel the mystique that has developed over the years that maps provide magical answers to our analytical problems (they do not), or are too complex conceptually (they are not) for everyday practitioners to understand. These misunderstandings about mapping are probably quite widespread. Understanding how maps are constructed helps in their application to data analysis.

When interpreting maps, you interpret data, but in a format that differs from standard analysis. A map does not create new information, but merely represents data and the relationships it contains. When thinking about maps, then, focus on the data to be analyzed in meeting the study's objectives. Some key issues that are critical to the interpretation of maps are:

- -The effect on maps and data analysis of including the correct set of objects/stimuli to be measured (e.g., brands, candidates defining the "market")
- -The completeness of the item list (e.g., the attributes used to rate brands)
- -The relation of the subject to the object/stimulus rated
- -Our criterion function (basis) for determining where we want to go or be with our brand, candidate, etc.

The simplest map is a one-item map of an individual item. We routinely present this type of data numerically, rather than spatially, although the latter can be a useful visual aid.

Figure 1 shows data that reflect consumers' perceptions of four spokespeople/endorsers of various products/product lines. Examples of endorsers are Jack Nicklaus or Cheryl Tiegs. An endorser conveys prestige and specialness to the products endorsed, although not necessarily through an experiential, direct association with the product category, such as Jack Nicklaus and golf balls.

Each of the four endorsers was measured on 14 items, using a scale that reflected how well the respondent felt the item described the person. These perceptions were tabulated and the differences were examined among the spokespeople to see how they were rated on each attribute.

The differences can be depicted in numeric form (e.g., average ratings) or spatial form as simple one-item maps. Figure 1 shows six of the 14 items; those that most differentiate the endorsers are apparent. Furthermore, the endorsers' relative positions on each are readily apparent, indicating which items best describe which endorsers.

Showing all 14 items in this way would give 14 one-item maps. However, such maps do little to clarify the data or summarize the results. We could choose to present just those items that showed statistical differences among the endorsers, but we would likely need to show too many. While one-item maps are of little practical value per se, they are the simplest way to show that perceptual maps reflect our original data.

Two-item maps do slightly better. The plot of the four spokespeople on the measures "well-recognized" and "I am familiar with" in Figure 2 shows an important general point: the higher a person scores on one item, the higher he scores on the other. In contrast, the two-item map of Figure 3 shows no such relationship among the measures. Whether a person is perceived as "adding value" is unrelated to whether he is "believable."

Figure 1

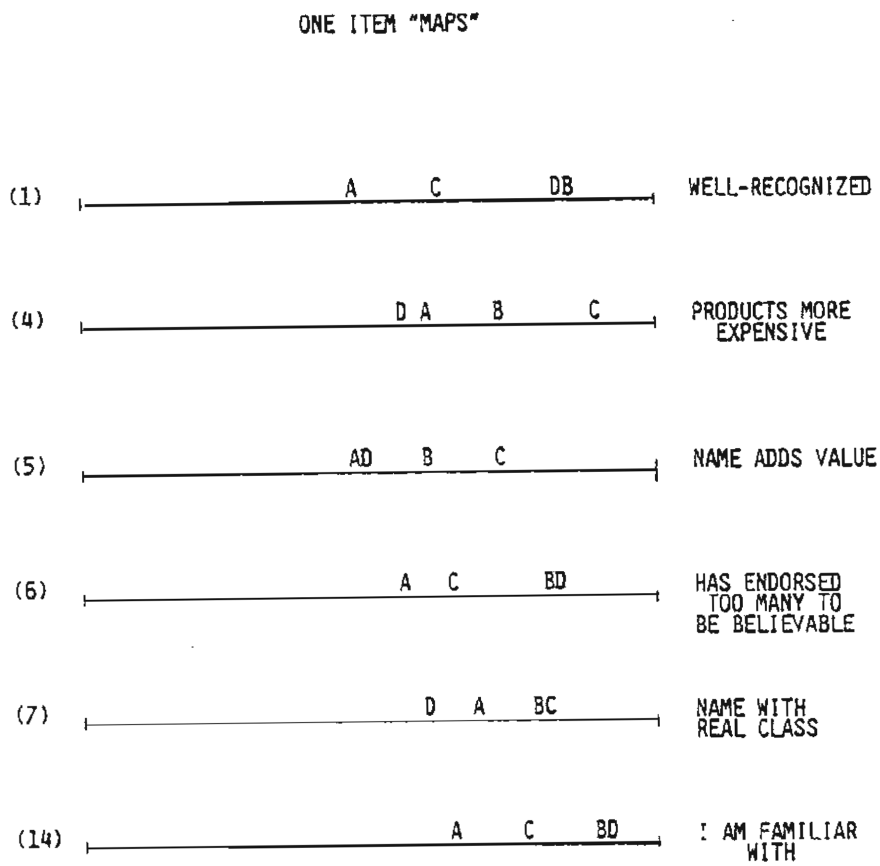
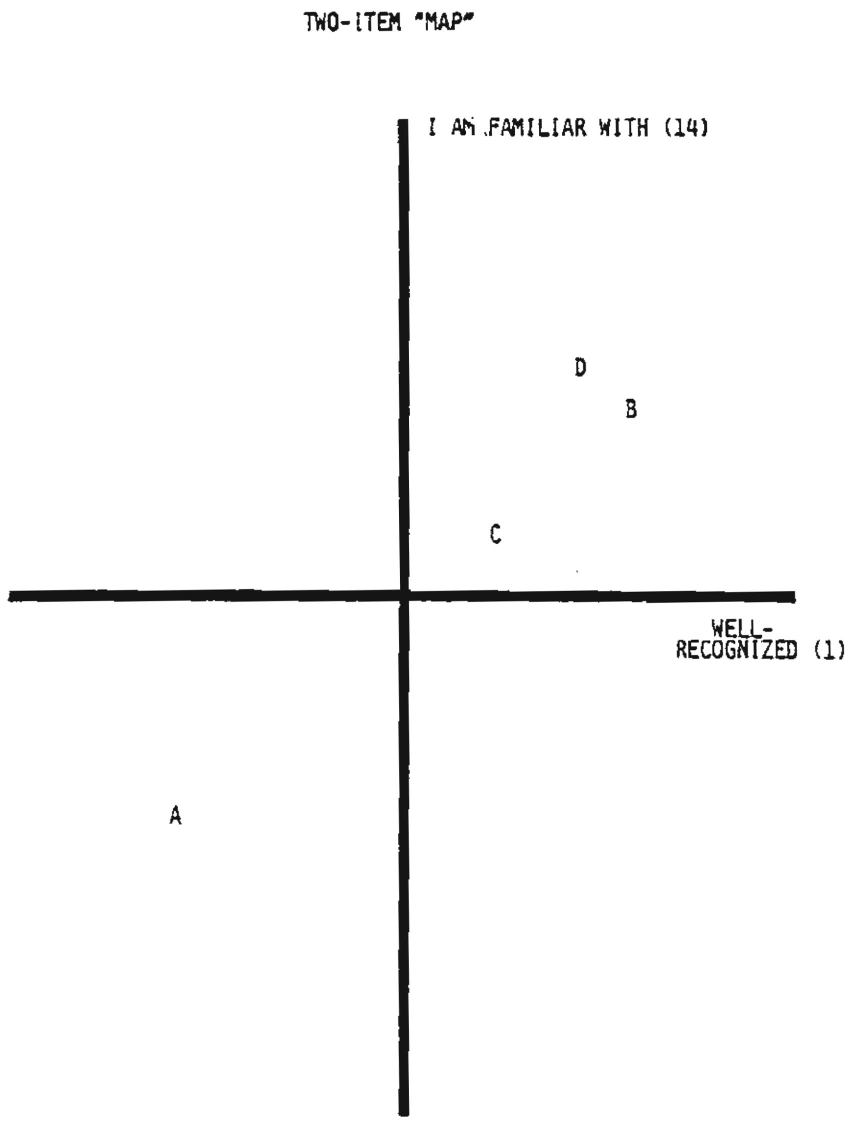
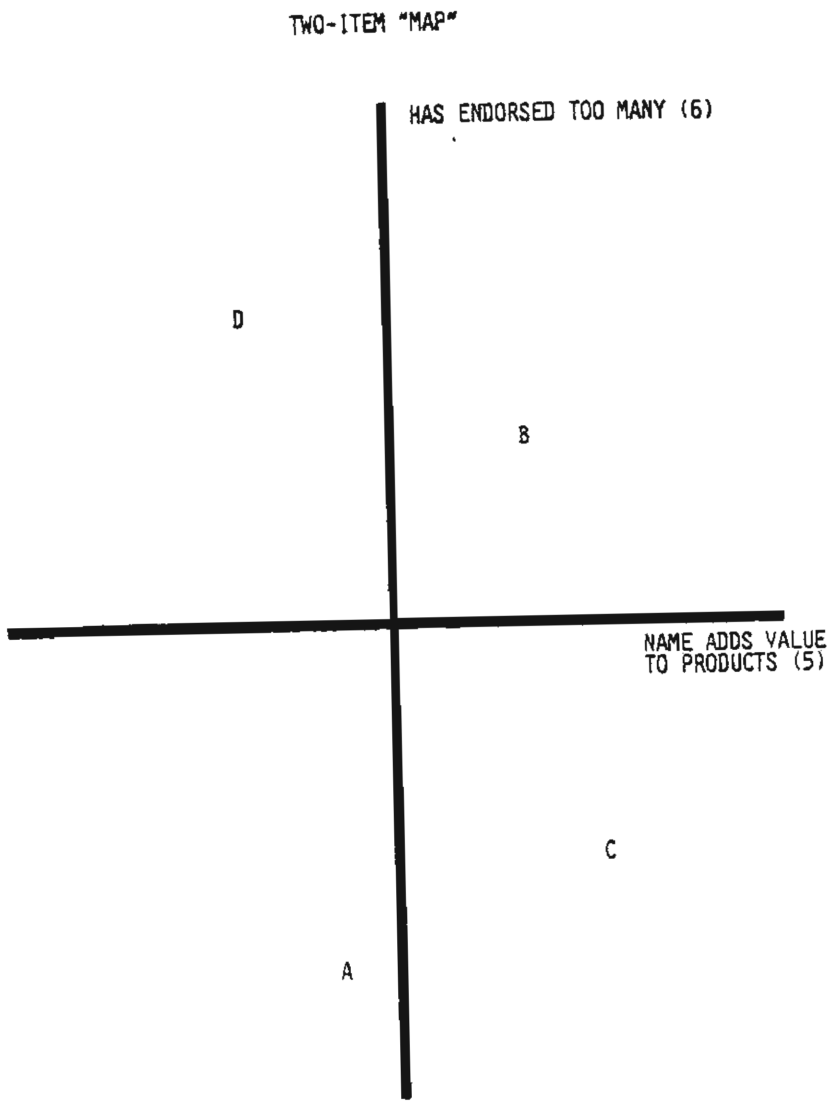


Figure 2



As with one-item maps, we gain little from two-item maps. While they do present results in a graphically pleasing way, we would need to draw 91 such maps, none of which help to summarize our results or draw conclusions.

Figure 3

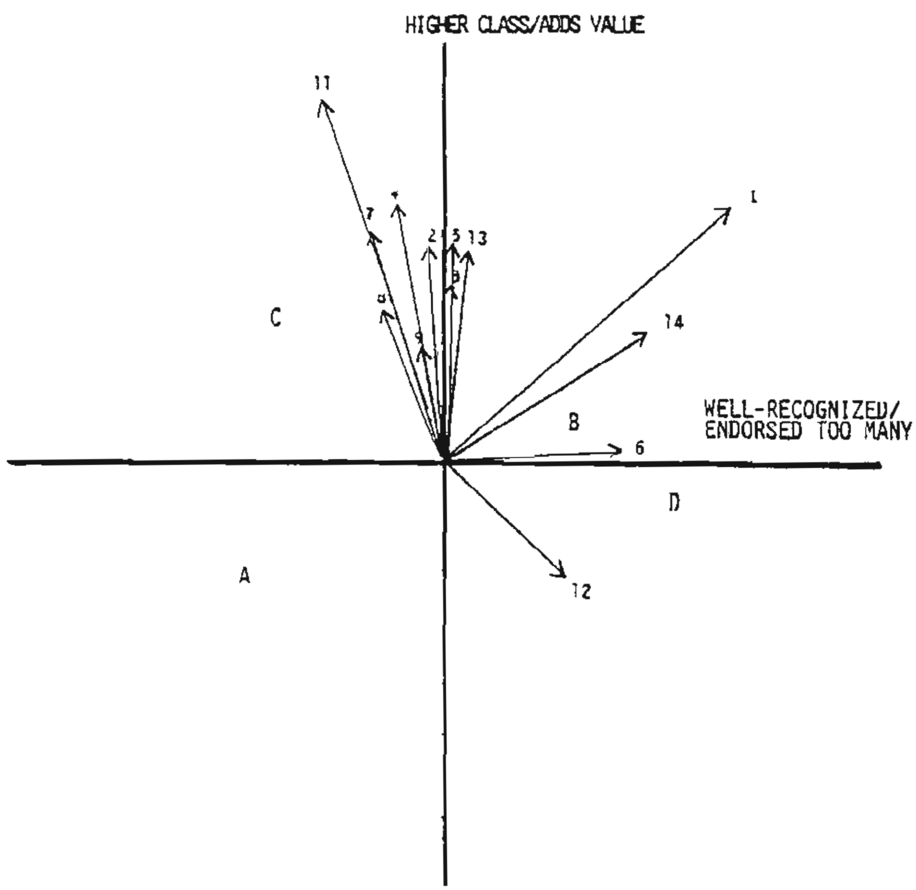


A map generated using discriminant function analysis (DFA) appears in Figure 4. The DFA algorithm seeks to find and combine items that most differentiate the objects being measured. Having done so, it seeks to find a second way of doing so, subject to the constraint that the second is mathematically independent of the first. Third, fourth, and higher combinations can also be formed, subject to certain limitations.

These combinations of items, or "dimensions," are comprised of two, three, or several of the original measures used. The original attributes that are combined in a dimension are often correlated, so that the net effect is data reduction. The number of dimensions used is smaller than the number of items, which reflects the "best way" to show the positions of the objects relative to each other (according to the criteria of "best" for the particular mapping algorithm).

The map in Figure 4 shows the four endorsers in the first two (most discriminating) dimensions derived from a DFA. The objects (endorsers) are represented by A,B,C, and D, and the 14 items on which they were measured by vectors or arrows. Rather than showing 14 one-item maps, or 91 maps of each possible combination of the two, we can show on one map the relationships among all the spokespeople on every measure.

Figure 4



The key questions are: How can we "capture" all the items in one map? How does the map reflect our original data? Once we have the map, how do we interpret the results?

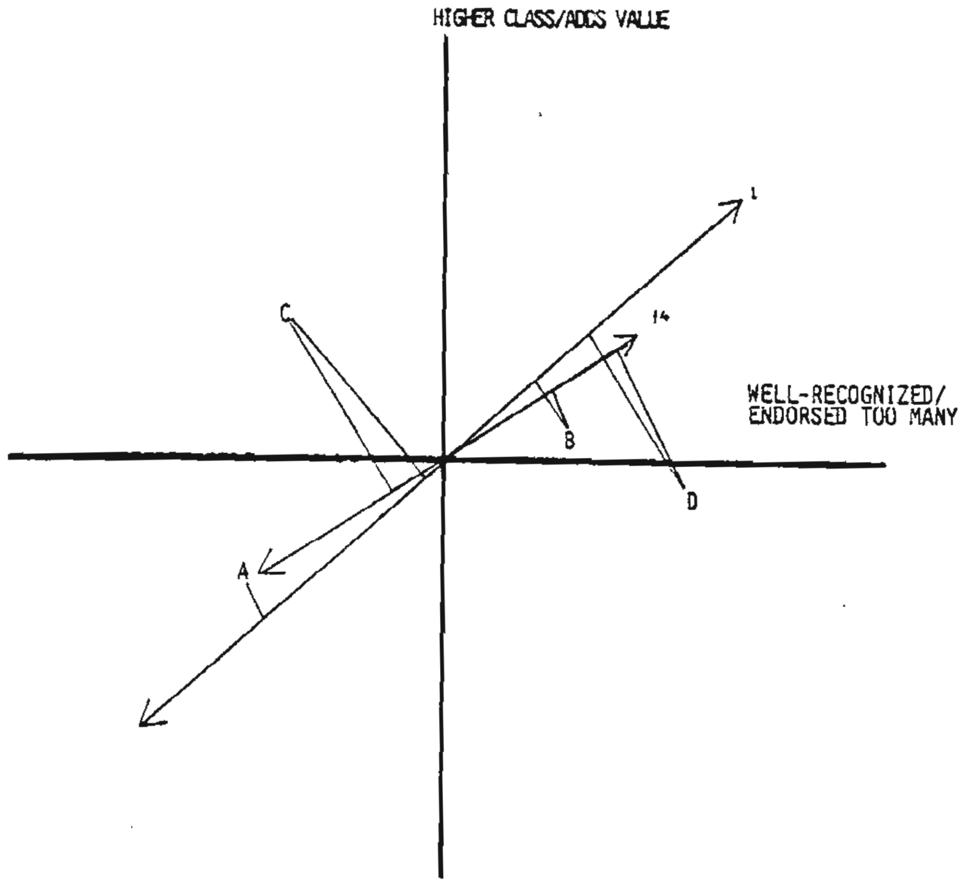
The answer to the first question is three-fold. First, although we measured our endorsers on 14 items, we do not have 14 pieces of information. Often there is redundancy in the items used, and the more the redundancy, the less we learn from overlapping questions. We can reduce our original set of measures to a smaller set, which allows us to describe our data in fewer dimensions. Since these dimensions comprise combinations of two or more individual items, they are "multidimensional" in nature.

Second, if the objects do not differ on some items, those items do not require dimensions on the map, especially if the criterion function to develop the map is based on finding how best to differentiate them. (Maps based on this criterion will mask items with "no differences," which can be an important omission from our data analyses. In marketing, it may be easier to establish a difference where none exists, than to modify an entrenched difference that is undesirable.)

Third, we are willing to give up some accuracy in order to gain simplicity of presentation - a tradeoff. When we summarize several items into a few, we lose something. Showing 14 items in a two-dimensional plot, even with similarity of meaning among some, cannot be done without giving up something. In addition to suppressing items for which there are no differences among the objects, dislocation of the relative positions of the objects on the items may occur.

Figure 5 demonstrates this point. It is the same as Figure 4, except all but two vectors have been removed. We relate an object to an attribute on a map by considering the vector for an attribute in the same way as a one-item map. We draw a perpendicular line from the object to the vector for the item of interest, which allows us to compare the relative positions of the objects on that measure.

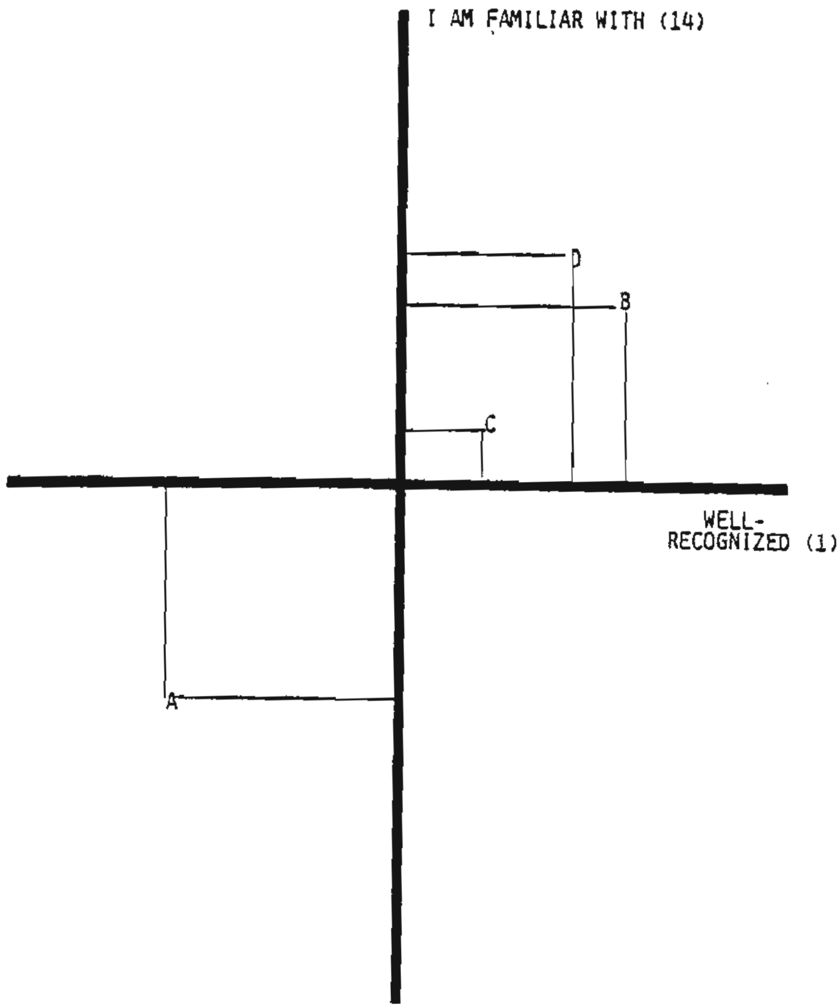
Figure 5



If we compare what happens when we make projections of the objects on the two vectors to the objects' relative positions in the original data, we do not obtain a perfect one-for-one correspondence (Figure 6). A reordering occurs on the "well-recognized" item from B-D-C-A to D-B-C-A. While the "error" is not great in this case, changes in item ordering like this occur and are a fact of life when attempting to summarize data from several attributes into a smaller set of dimensions. The creation of a map may require us to accept some dislocation in the relative positions of the objects. As emphasized earlier, the map, no matter what type, is merely a way to represent the original data.

Figure 6

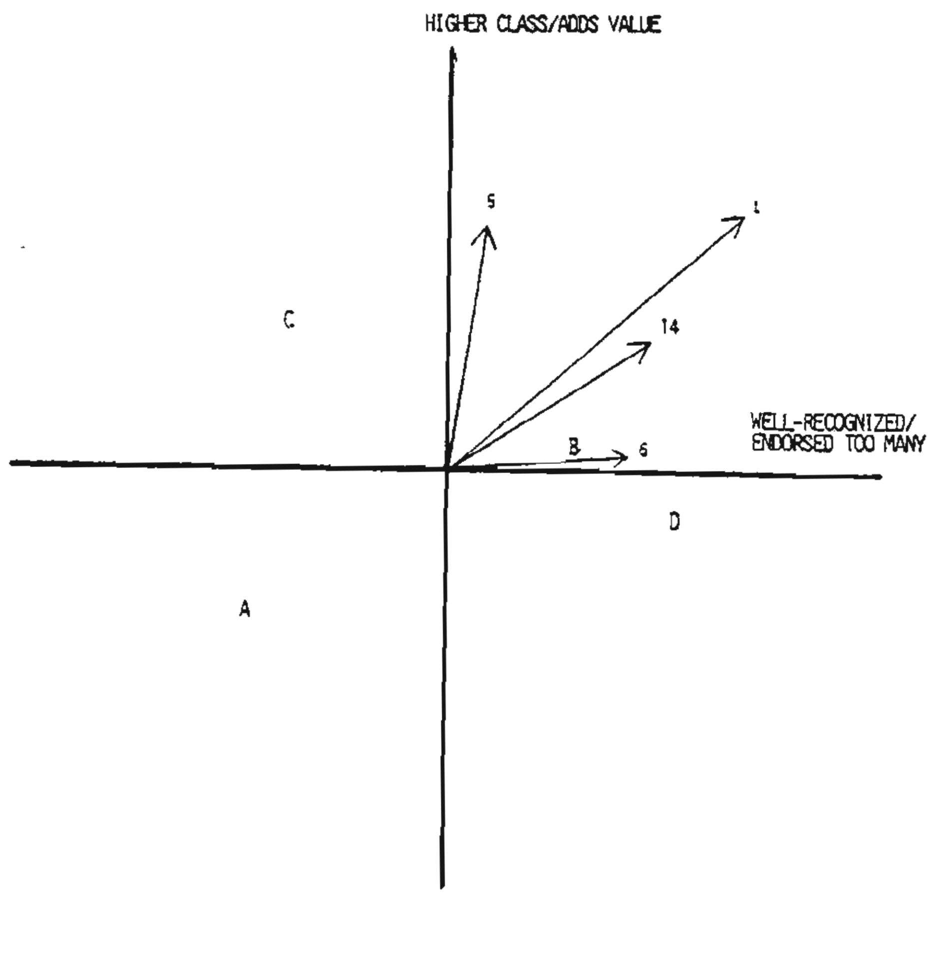
TWO-ITEM "MAP"



Once the original data are represented in a map, the focus turns to their interpretation. The first step is to name or identify the dimensions we created by combining sets of the original items, a process directly analogous to naming factors in factor analysis.

To do this, look for vectors that are long (length represents the correlation between the item and dimension), and which lie close to one of the derived (independent) dimensions. While several items qualify, I have chosen item 5 ("adds value") and item 6 ("has endorsed too many to be believable") as representative of how to name the dimensions in Figure 7. Items 1 and 14 are at 45 degree angles to the two dimensions and lie close to each other. This means that they are associated with both dimensions.

Figure 7

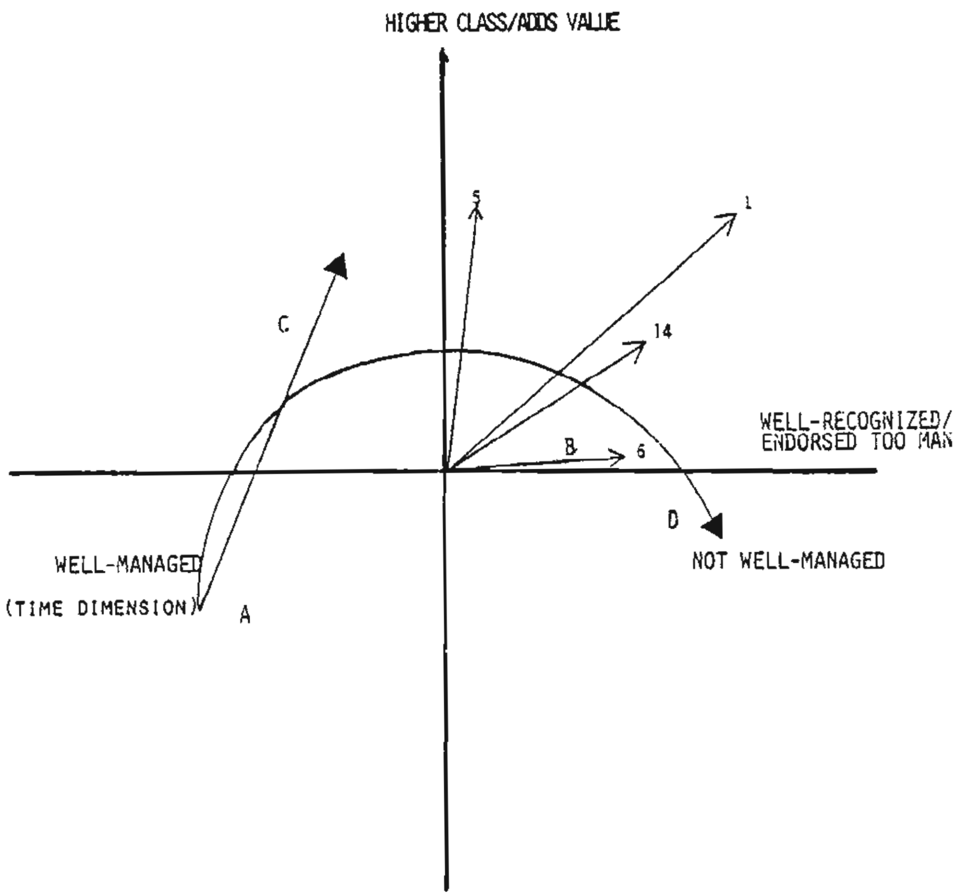


Based upon this set of names for the dimensions, endorser D is positioned unfavorably, and endorser B is not strongly positioned either. On the other hand, endorser C is strongly positioned, being perceived as adding value and not having diluted the value of his/her name. Endorser A has neither value nor recognition.

I have pointed out that maps are merely a reflection of the original data and contain no magical qualities. However, maps can be helpful. The visual representation can suggest an interpretation that likely would not be made from a numeric presentation.

The ultimate goal in examining maps is to derive an interpretation that suggests not just where each endorser is positioned, but also where we want him/her to be. The point of view we developed, assisted by our knowledge of who these spokespeople are, is clear from the time/management dimension we overlaid on the map (Figure 8). Spokesperson C has struck

Figure 8



a sound balance between exposure and value, and will most likely experience continued success if the same management philosophy is pursued. Further, endorser A has the opportunity to be developed along the lines of C, but B and D have gone too far and should maximize their efforts to maintain their current product lines.

The interpretation above was based on all respondents. Further interpretation is possible based on conditional maps, which show how a map can be affected by the relationship of the respondent to the object. This relationship is usually cast in terms of awareness, knowledge, and usage of brands of products and services.

In the interpretation of maps, the relationship between respondent and object can have a significant effect on recommendations and marketing actions. Considering the factor of awareness, with four endorsers there can be 16 possible combinations of recognition/no recognition among them. More endorsers would mean more possibilities. Usually, a small subset of awareness combinations captures the bulk of the market. Even if they did not, we would not want to look at all of them.

We must take conditional maps into account, because the ramifications can be significant. As an example, consider two extreme cases: Those with higher awareness of all four endorsers and those with lower awareness (Figures 9 and 10). As shown in the maps, the positions of the endorsers differ by condition, as well as from the map of all respondents. In particular, among those more familiar with all, endorser D is in a worse position than among those less familiar. As the awareness set for objects changes, it is important that our analysis accommodate this dynamic aspect. In marketing, it is far easier to establish a desired image among those unfamiliar with our product or service, than it is to modify a strongly-held undesirable one.

Figure 9

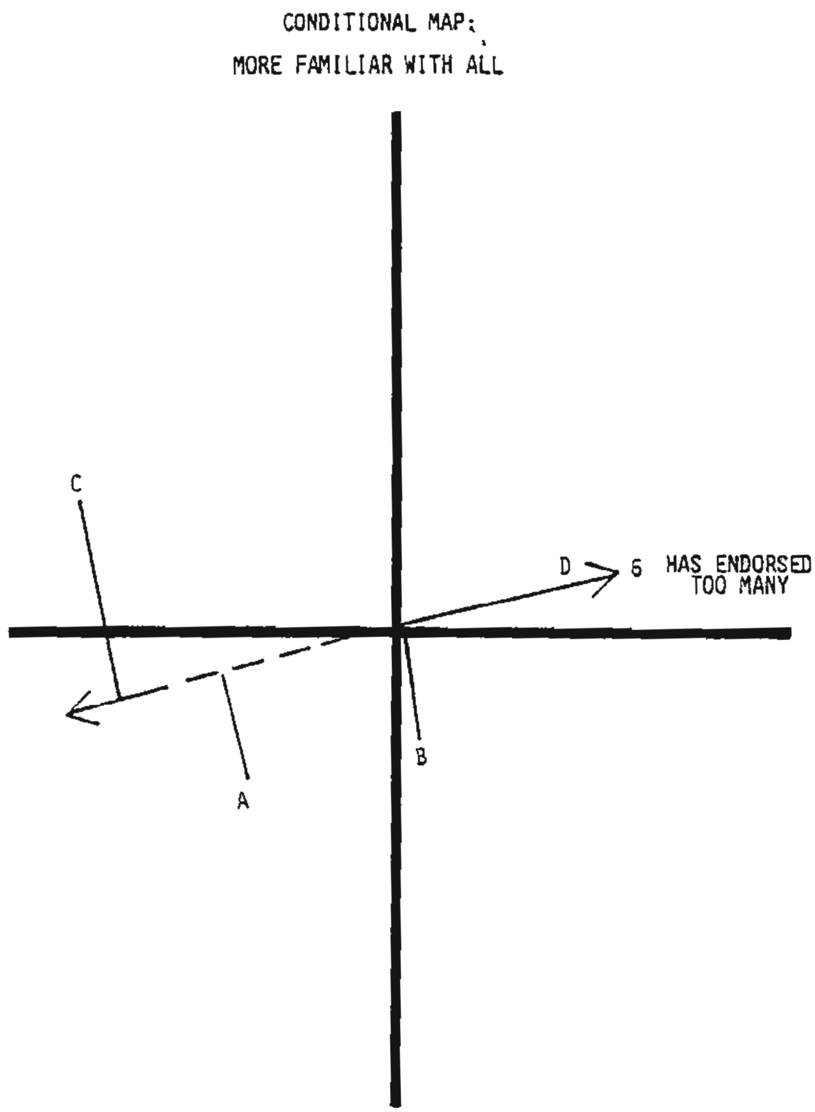
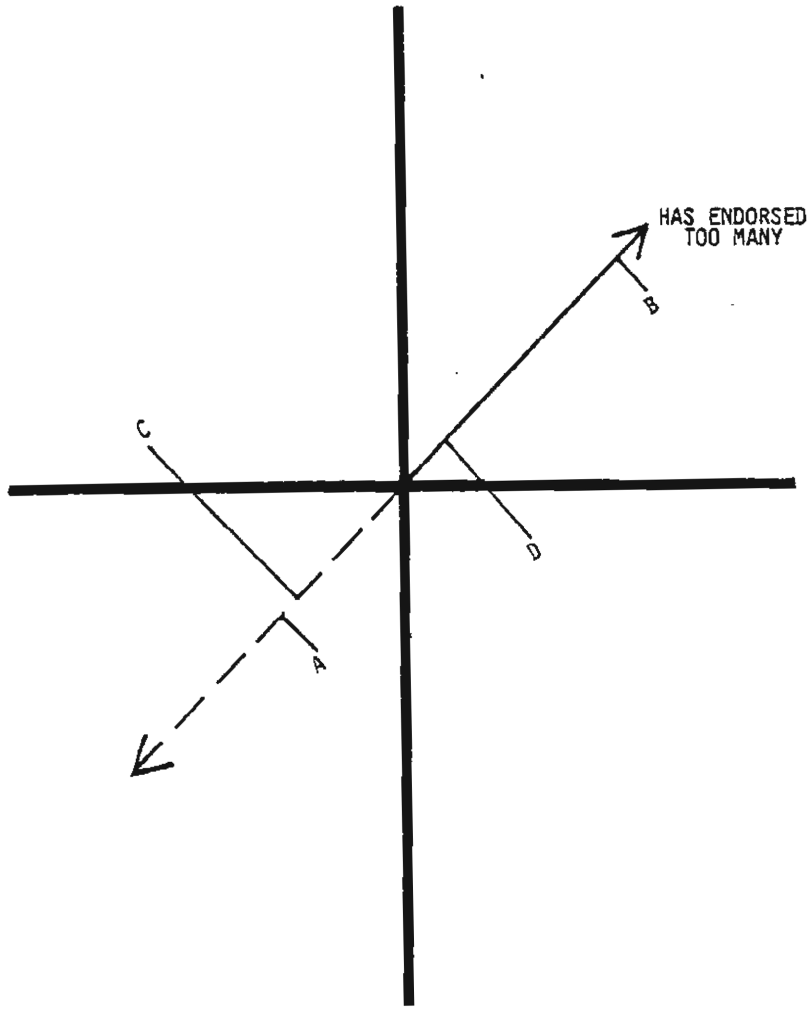

CONDITIONAL MAP:
MORE FAMILIAR WITH ALL
C
D
6
HAS ENDORSED
TOO MANY
B
A

Figure 10

CONDITIONAL MAP:
LESS FAMILIAR WITH ALL



In addition to paying attention to the relationship of the respondent to the objects being evaluated, it is also crucial to correctly define the relevant market - the set of objects that we compare to one another. This point is illustrated in Figure 11, which reflects consumers' perceptions of large national financial service institutions. We obtained a DFA map that shows separation on the two dimensions of their willingness to customize their products, and the broadness of their expertise and product line. Based upon this map, ways to modify the product to better compete could easily be developed.

However, by adding just one additional competitor ("B"), who lacks national scope, broad expertise, or ability to customize, a new most-discriminating dimension emerges (Figure 12), which did not appear in Figure 11. The omission of this competitor from the first analysis is crucial because he has a high share of most consumers' financial business, and he is trying to broaden his financial activities. All major players in a market must be included in the data analysis, otherwise the conclusions will miss the mark.

Figure 11

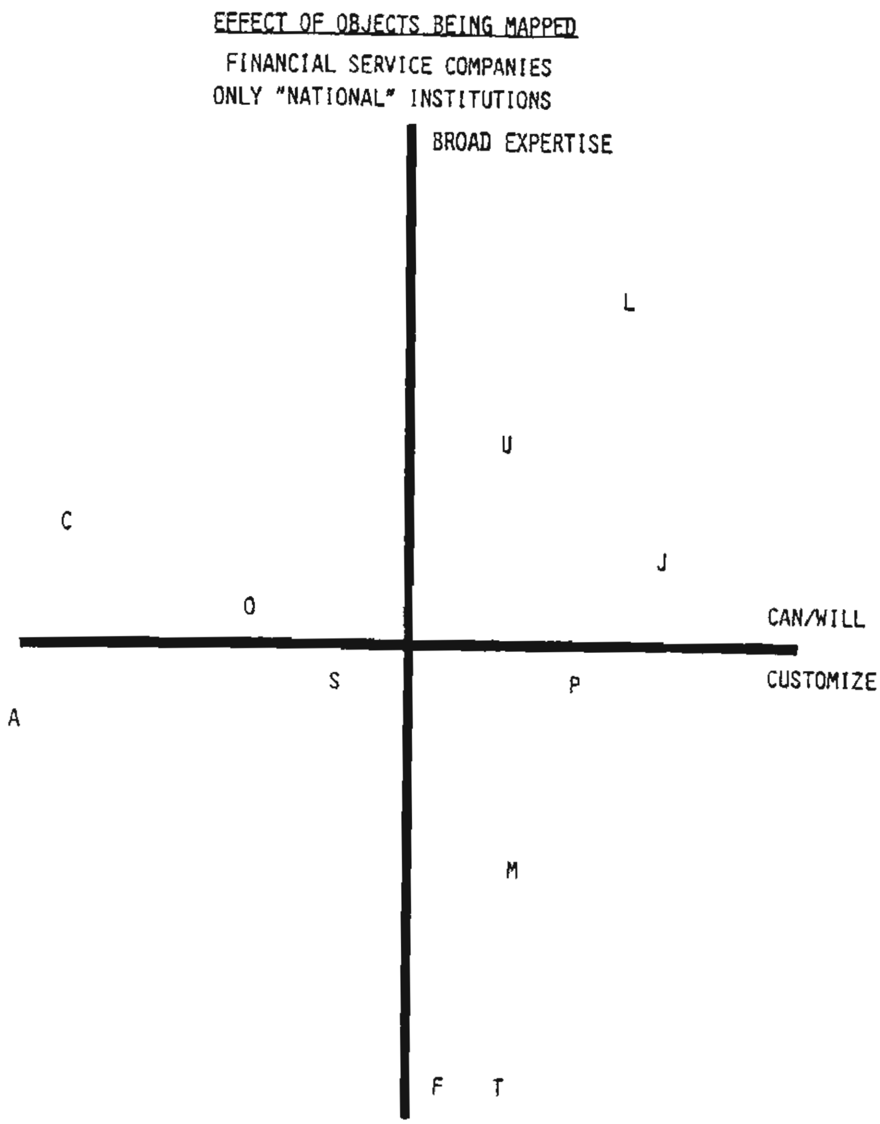
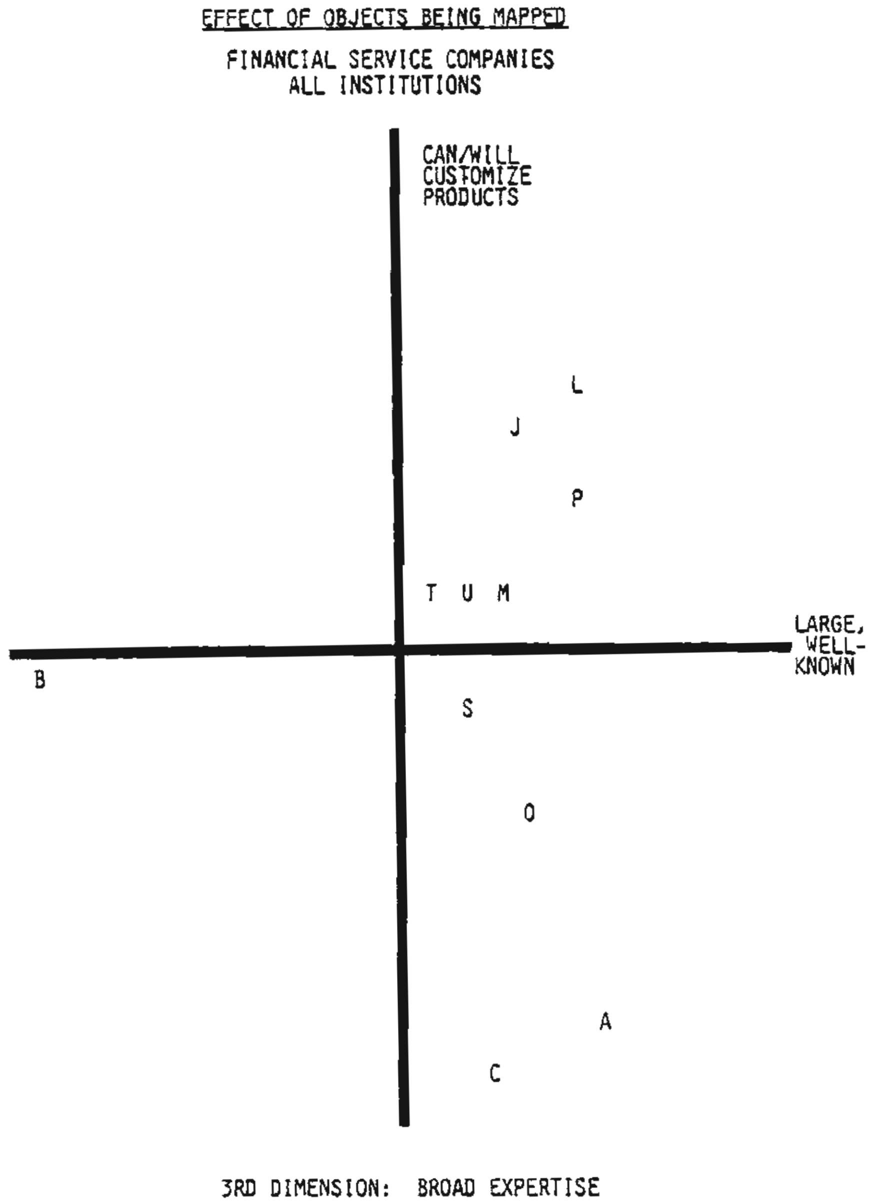
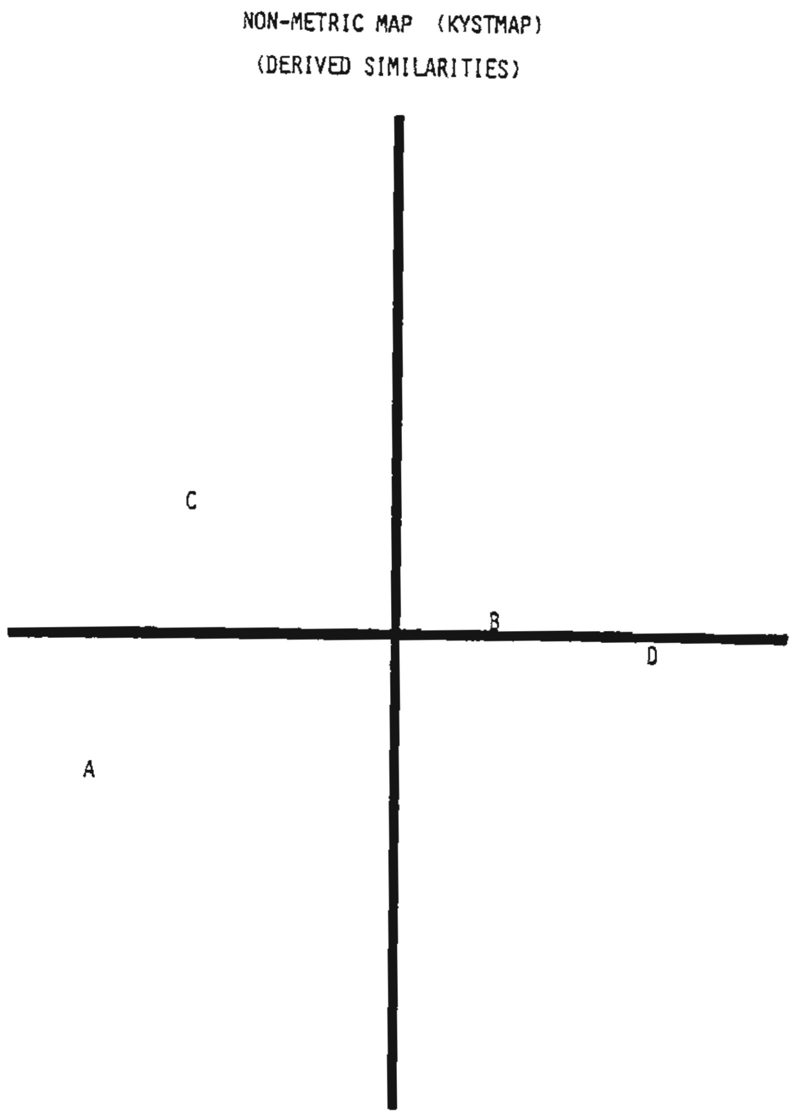


Figure 12



One "technical" topic with pragmatic ramifications is the type of measure and mapping program used to develop a map. To demonstrate this point, I subjected my endorser's data to the nonmetric scaling program called KYSTMAP. This was done by first deriving similarities measures among the endorsers using each respondent's Euclidean distance for all possible pairs. These nonmetric data were then analyzed by KYSTMAP, which, because of the nature of the derived similarities data, had no prior knowledge of the items that were used to generate the distances. Whether the same result would have occurred with some direct measure of similarity (as opposed to metric measures) is unknown. However, Figure 13 shows that the KYSTMAP map is very similar to the one generated by DFA.

Figure 13



The debate is not which program or scaling method to use, but rather which facilitates analysis and is more likely to give valuable direction. If one uses direct measures of similarity or dissimilarity in the absence of any metric measures, the analyst must interpret the map. Without the availability of metric measures, the KYSTMAP map may not have led to the development of a sound theory about each endorser's respective position. This difficulty of analysis is not uncommon, and many who support nonmetric measurement administer batteries of attribute questions to help in their analysis.

Nonmetric measurement often results in consumers revealing just the most obvious in their similarities judgments, such as dimensions relating to frozen vs. canned packaging, high vs. low price, or high vs. low calorie. Thus, little or no insight is gained.

In conclusion, if perceptual maps are one of the modes of analysis, use metric maps derived directly from metric measures. There is everything to gain in ease and richness of analysis, and nothing to lose. Some believe that direct measures of similarities might uncover an issue or attribute that the marketer/analyst had overlooked; having a complete set of items is crucial to the success of any study. I have never seen a map uncover a "new" dimension. Instead, a clear understanding of the market and the proper development of item lists prior to fielding studies provide all the measures needed for a complete analysis of a market.

I have discussed various pragmatic issues concerning the interpretation of maps; some simplistic notions of their development and their relation to the original data; and one technical point. One serious shortcoming has been evident in the maps presented: They have all been static, representing a fixed set of positions of objects in their respective markets. We MUST strive to develop and implement solutions to this serious shortcoming in order to increase the power of our data analyses.

The term "dynamic" when applied to conditional maps means that consumers move from state to state. In order to understand a market we must take this dynamism into account. There are many types of dynamism, reflecting the fact that most markets are extremely complex environments with several possible problems:

- Not all people are aware of/have knowledge of the same products/services

- Not all people define a "market" in the same way

- Different people attribute different sets of characteristics to products/services

- Different people buy the same product for different reasons

- Different people buy different products for the same reason

- Communications of marketing actions do not reach all the consumers intended

As marketers we attempt to affect these interrelationships. We must strive to develop a more comprehensive analytical model that accounts for all these complexities, and which allows us to evaluate all kinds of marketing strategies.

There have been many models of this type described in the literature, and some marketing research services offer models that include the capabilities listed. Unfortunately, they have not reached widespread understanding, use, or acceptance. Most analyses are still static. To increase our effectiveness, we must begin to apply more powerful models that allow us to address the kinds of questions management needs answered. Note that my operative word is "model" not "map." We need to focus on models of markets, which can take on many forms predicated on different kinds of data.

What then is a map? Maps are very useful adjuncts to the types of models just described, which provide insights into our data. However, maps must not be communicated as ends unto themselves. They are vehicles of analysis, helping to take us to a higher level of understanding, beyond statics into dynamics.

I hope I have successfully put maps into their proper perspective, while still conveying their value. If so, I can revise my original set of alternative titles to reflect a more actionable role for maps in data analysis:

- Vectors, Vectors Everywhere, and They're Telling Us Where to Go

- The Elusive Market Niche- I Can Find It (Them)

- The Fifth Dimension: The Competitive Edge

- We Are Doing More than Map!

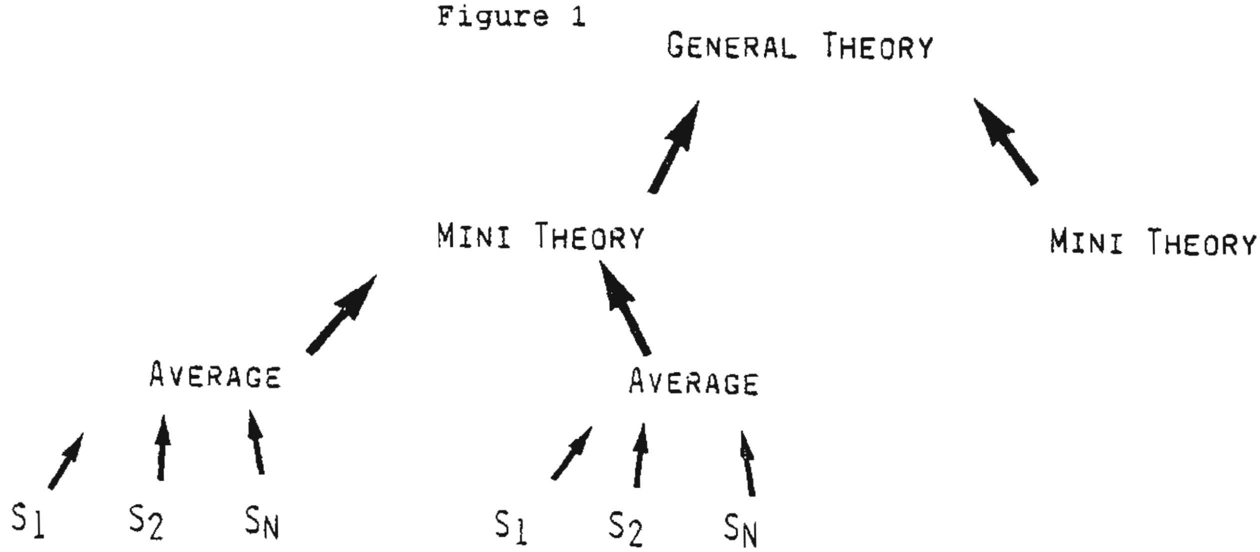
PRESENTING RESULTS

by
Bruce Morrison
General Electric
Louisville, KY

In the introduction to his 1972 book, Multidimensional Scaling Theory and Application, Roger Shepard states, "The unifying purpose of [the many techniques loosely subsumed under the term "Multidimensional Scaling"] is to (A) somehow get hold of whatever pattern or structure may be hidden in a matrix of empirical data, and (B) represent that structure in a form that is more accessible to the human eye - namely, as a geometrical model or picture."

In essence, this is the basis of science - reduction of information, rather than the generation of information, as some would suspect. The reduction occurs from raw data with complete information on each individual to averages and correlations that are fictional. These give rise to mini-theories that incorporate the results of this study and similar studies, which may evolve into a general theory of human behavior (Figure 1).

Figure 1



Why does science undertake this reduction of information? First, to be able to generalize to new areas. For instance, information on a particular airplane could lead to a hypothesis on how a proposed airplane would react. Second, and more important, reduction makes information easier to use, remember, manipulate in the mind, and apply to everyday situations.

The major point of my presentation is that if the information we generate is not used, it is by definition USELESS, regardless of how inventive or good it seems. Using this operational definition, I was depressed to find that many studies I had previously considered my best were useless, whereas others I had considered light and frivolous had stood the test of time and are still being used and presented 15 years later.

What made the difference? The presentation. These presentations were to the "right" audiences, and regardless of how complex the data, were presented in a form that could be remembered by the audience for daily business use.

There are several audiences we make presentations to, each with separate underlying motivations. They include research directors, brand or product management, upper management, and sometimes the "movers and shakers." We are all familiar with the first three, but who are the movers and the shakers? They are the individuals who can move the program forward or shake it to the ground. For instance, in an advertising agency they are the lowest art directors and copy writers who actually do the work. In a company, they are the salesmen who sell the products and execute the plans. In publishing, they are the people who advertise and pay the bills. In trying to affect the general public, they are the news writers, radio talk show hosts, and television newscasters.

Knowing their research understanding and motivation helps us to understand each of our potential audiences (Figure 2). In terms of motivation, market research directors are high, because they have to understand, interpret, present and justify the expense of the study. Brand/product management are generally low, because they are not responsible and do not have a vested interest. Any change means more work for them. Upper management are midway, due to the financial responsibility of new products, advertising and programs. Movers and shakers are low, because for them it only means more work or changes in procedures and/or new criteria.

Figure 2

	RESEARCH UNDERSTANDING	NEED TO UNDERSTAND "MOTIVATION"
MARKET RESEARCH DIRECTORS	HI	HI
BRAND/PRODUCT MANAGEMENT	MID	LOW/MID
UPPER MANAGEMENT	LOW	MID
MOVERS & SHAKERS	LOW	LOW

If you don't affect the market research director, you stand to lose potential projects. If you don't affect brand/product management or upper management, you have a serious chance of losing your credibility as a "nuts-and-bolts/action-oriented" researcher. If you don't affect the movers and shakers, nothing gets done.

The program or product will be initiated, but not with the enthusiasm and effort which will ensure its success. But how do you present to this audience of individuals, who regardless of their expertise in other areas, have little understanding of marketing research or motivation? In the words of some long-forgotten advertising genius, the answer is "K.I.S.S.- KEEP IT SIMPLE STUPID."

We have our new techniques and methods, our six-pen plotter that can produce three-dimensional graphs, but the only thing your important audience wants to know is: What does it mean for ME? Marketing research should produce answers, not dazzle with its footwork. This does not mean you should avoid impressing the audience with your brilliance, but get on directly to their concerns.

How do you stage a good presentation? First, a good presentation begins with the design of the study. In addition to the head of marketing research and the vice president of marketing, talk to the movers and shakers. Many times they are the grumpy old salesmen of whom even the vice president of marketing is afraid. They will tell you that there's no need for research. But if you listen closely, you will find contradictions and questions. If your presentation answers these questions, everything else you say will be taken as gospel.

Second, once the data have been analyzed start to "telegraph" the findings to the key individuals with the most to lose. Any study that shows individual "A" was wrong in his decision will be attacked by individual "A," unless he has had a chance to think about it and "get his ducks in a row."

Third, make the presentation memorable. When I did heuristic clustering of how people perceive sex in advertising, I was able to publish an article in the Journal of Advertising Research (which has never been read by anyone I know). When I took the same information to the Ad Age Creative Workshop and talked about "Tom Jones" - if it moved, fondle it - I got 53 column inches in the "Los Angeles Times," a reprint in Reader's Digest, and television and radio interviews that have continued for 15 years.

When in 1973 I used an INSCAL analysis, PREFMAP, and psychographic clustering of industrial perception of advertising, my Journal of Marketing Research article was accepted with revisions. When I was able to generate word pictures that were translated into cartoon characters, the client produced a 23-minute, fully-animated cartoon that has been shown to over 235 companies in the New York City area alone. Why? Because I did what multidimensional scaling was supposed to do, according to Roger Shepard in 1972. I turned a matrix of data into something that is more accessible to the human eye and memory.

This all sounds great, but I have a few cautions and concerns to advance: First, expect to be played up as a genius far above your wildest dreams. At the same time, expect to be made fun of as the wild-eyed, long-haired genius. Second, work directly and personally with those who "help" you with this process. This helps protect the integrity of the data, which may be simplified to the extent that you may not recognize it.

As our technical ability increases, we will more and more have this problem of how to present our findings in a way that is understandable and memorable. Men and monkeys can be creative because they rearrange and apply mental concepts. If data cannot be remembered by the person who makes the decisions without referring to a file drawer or the research director, the data are not being used to their fullest capability.

PRESENTING THE RESULTS OF PERCEPTUAL MAPPING STUDIES

by
John A. Fiedler
POPULUS, Inc.
Greenwich, CT

I will focus on the presentation of maps, once a strategic study has been designed and sold for which perceptual mapping is a key component. The study has been fielded, and the data analyzed with the Sawtooth Software APM System.

The purpose of including perceptual mapping as an analytic tool is ultimately to present the data in a map format. A map is a marvelous presentation tool that contains an enormous amount of data that can be readily absorbed by your audience. I will discuss five rules related to map presentations and illustrate each with data collected by the APM System.

Rule #1 - Make It Interesting

As the researcher, make certain your audience is interested and involved with your presentation. Much of what Ted Evans said previously about selling within an organization is critical to enhancing interest. Another key element of selling within is the participation of the audience with the questionnaire task itself. Have them sit before a monitor and complete the questionnaire before the presentation. Many are more interested in a computer study because the interviewing task itself is more interesting. Your audience should be genuinely excited about the forthcoming results.

I would have liked to have presented the findings of our clients' studies, but the maps belong to them. Instead, I sent the 40 speakers on this program a diskette with a questionnaire about the character traits of various Presidential candidates and the issues shaping the 1988 campaign. Much to my surprise, 80% of the diskettes were returned. Several others in the speakers' organizations also completed the questionnaire, yielding a final sample of 135. I assume this response was motivated by the promise of an interesting presentation. One portion of the data about Presidential character traits will be discussed in two dimensions, which brings us to the second rule.

Rule #2 - Present Only One Map

"One map" means the quadrant space based on just two dimensions. This is all that should be necessary, since an enormous amount of information can be contained in this space. Allan Shocker showed how much information about breath fresheners can be contained in one product space. Moreover, it is enormously confusing to define more than two dimensions in the same presentation. It's a rare situation when information cannot be reduced to two dimensions during a management presentation meeting (not the research meetings beforehand). Your analytic task has not been done thoroughly if the data cannot be distilled to two dimensions for the presentation.

One of the advantages of the APM System is that it lets you try countless dimensions and rotations, allowing the investigator to do all the work before meeting with management, rather than forcing the audience to do this work for the investigator.

Rule #3 - Present the Map Very Slowly

The most critical rule to follow in presenting a perceptual map is the "Gypsy Rose Lee Rule." Present it very slowly, revealing one interesting piece of information at a time. This insures that the audience will follow your logic, and you can build or tell a story with increasing interest.

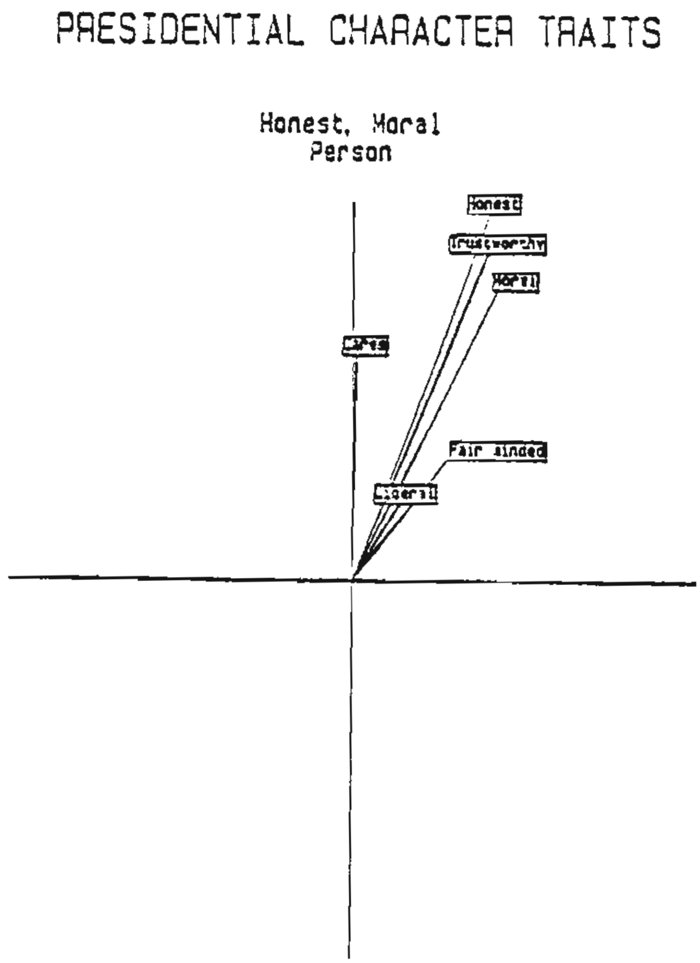
Present each dimension and the attributes that define that dimension first, followed by products. Use this order because the audience is much more interested in the products than the dimensions or attributes. At the end of the presentation, the focus of the discussion will be on those product points on the map. After all, strategic issues were the objectives of the perceptual mapping study: Should we move the brand to a different place? What if Brand X moves? and so on.

First, set forth a clear understanding of each dimension. Rather than presenting vectors on the screen simultaneously, present and explain the dimensions one at a time, using only those vectors that are highly correlated to that particular dimension and ignoring the unrelated vectors.

The data from this particular sample of 135 may not be typical of the American People as a whole. While many in the audience who completed the questionnaire view Jesse Jackson as a mainstream candidate, most Americans do not!

After some selections, rotations, and a great deal of analysis, the first dimension emerged called "Honest, Moral Person," which is described by such items as "cares and concerned about people," "honest," "trustworthy," and "moral." These attributes are all highly correlated with this dimension and uncorrelated with the second dimension (Figure 1).

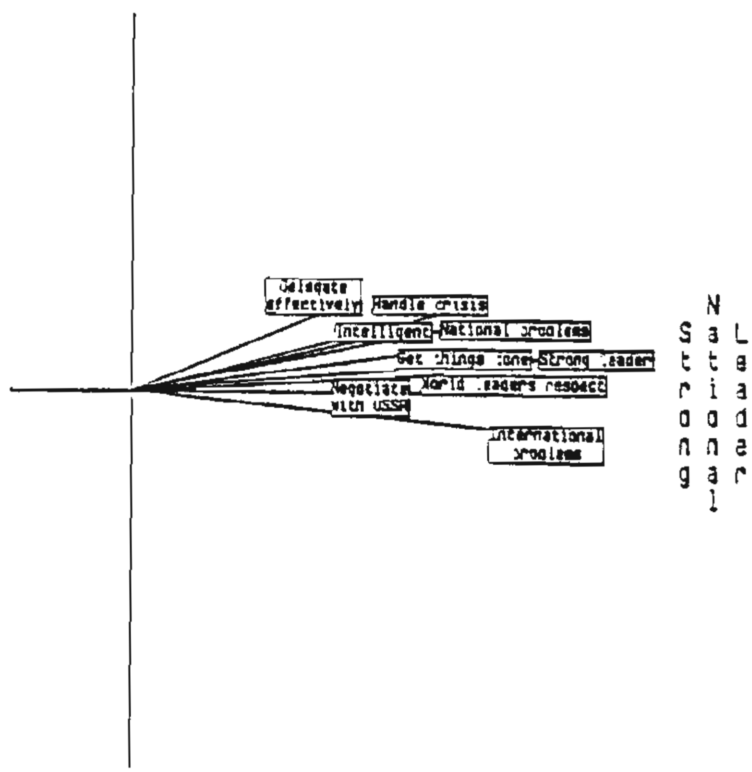
Figure 1



"Strong National Leader" is the second dimension, which consists of such items as "knows and understands a lot about international problems," "negotiating an arms control agreement with the USSR," "is respected by world leaders," "gets things done," "knows and understands a lot about national problems," "can handle a crisis," "is intelligent," and "delegates effectively" (Figure 2). In a presentation, take as much time as necessary to explain the attributes and why they are grouped with the label of the dimension.

Figure 2

PRESIDENTIAL CHARACTER TRAITS



The items "fair-minded" and "liberal" are not correlated with either dimension. "Liberal" and "conservative" formed a dimension of their own. While this is interesting methodologically, it is rather predictable and, as a dimension, boring. This kind of outcome is best presented in a table of means and is not worth the graphic space.

Once the two dimensions have been presented, it may be worthwhile to summarize them for your audience. In this case, one dimension is an evaluation of the "goodness," "worthiness," or "morality" of Presidential candidates. The other independent dimension shows how they are perceived in managing the role and Office of the Presidency.

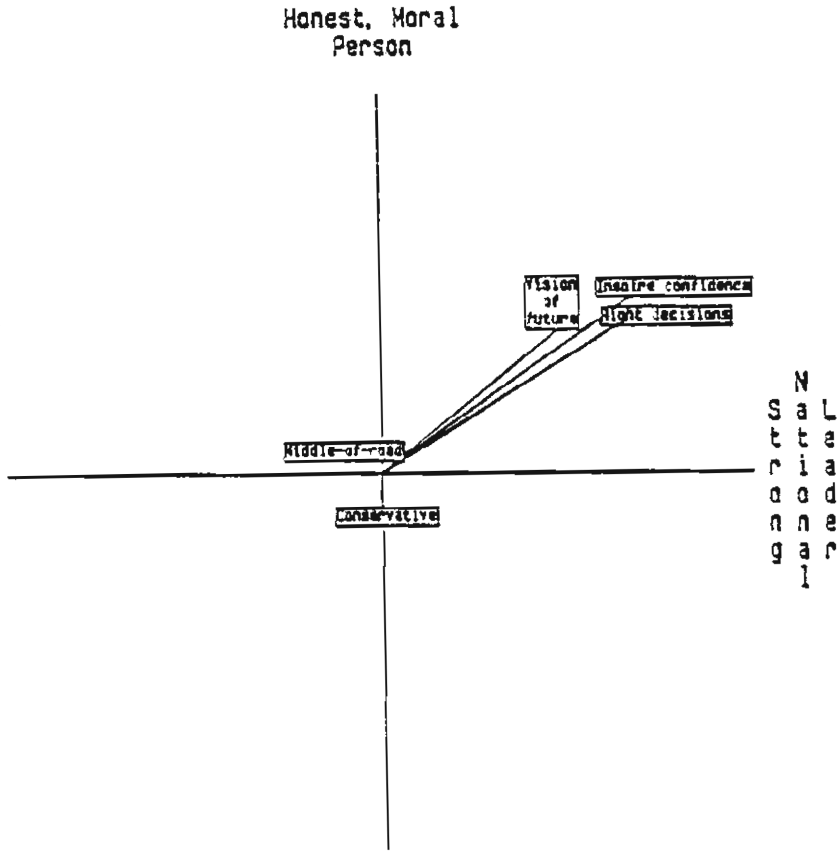
This is a good time to explain the items that don't fit the axes. In this study there were few uncorrelated elements. "Middle of the road" was unrelated to anything, which demonstrated face validity. Finally, some items such as "has a strong vision of the future," "inspires confidence," and "can always make the right decisions" gave 45 degree angles (Figure 3).

These three items combine aspects of both leadership and morality. A candidate perceived in the top-right corner of the space is like Plato's "Philosopher King." Interestingly, in a factor analysis these items fall out as a fourth dimension. Clusters of vectors not exclusively correlated with one dimension or the other can be presented and explained as these items have been. One of the advantages of using discriminant analysis rather than factor analysis is that items that form a separate factor in a factor analysis often become a cluster of intercorrelated attributes in a discriminant analysis.

At this point in the presentation, you can put aside all the vector maps. People understand what it means to be higher up on the map ("Honest, Moral Person") and towards the right ("Strong, National Leader"). Both axes with their labels "Honest, Moral Person" and "Strong, National Leader" can be shown on one map.

Figure 3

PRESIDENTIAL CHARACTER TRAITS



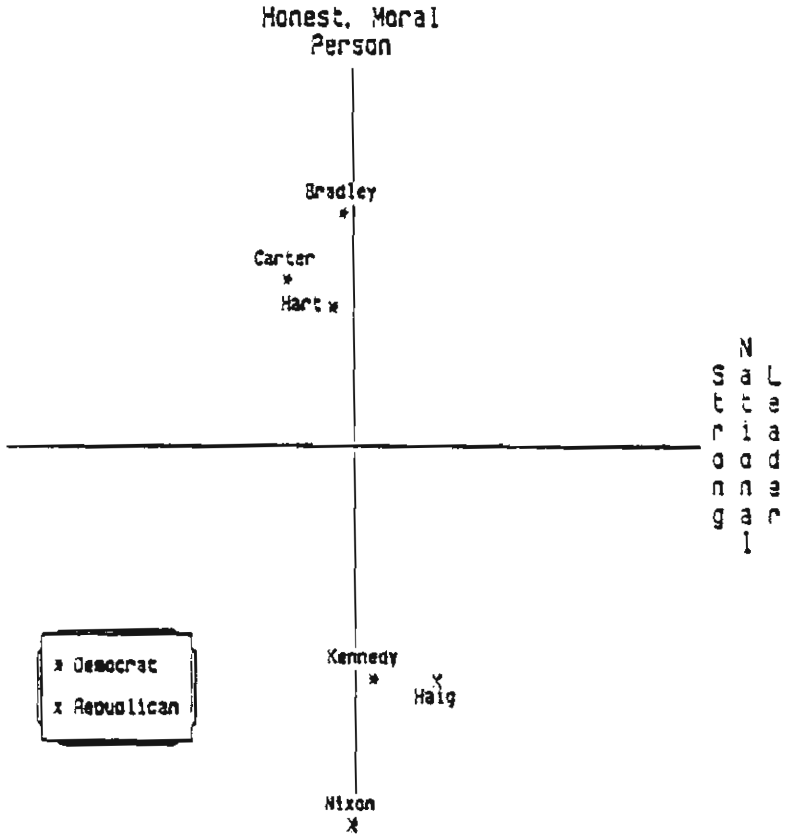
Rule #4 - Present the Products One Or A Few At A Time

You can control the development of people's thinking by how you orient them to the map, thus reflecting your chain of thought. By this time, any audience is eager to know which candidates characterize this "Honest, Moral Person" dimension.

Figure 4 shows that Bill Bradley, Jimmy Carter, and Gary Hart were seen as being highly moral and honest. (The disclosures that led to Hart's withdrawal of his 1988 Presidential bid occurred subsequently.) On the negative side emerged Richard Nixon, one of the past four Presidents included in this study, along with Al Haig and Senator Edward Kennedy. Some people might say that Kennedy solved the "dilemma" of Democrats clustering near the top of this dimension and Republicans at the bottom, but the perception likely is attributable to his Chappaquidick heritage.

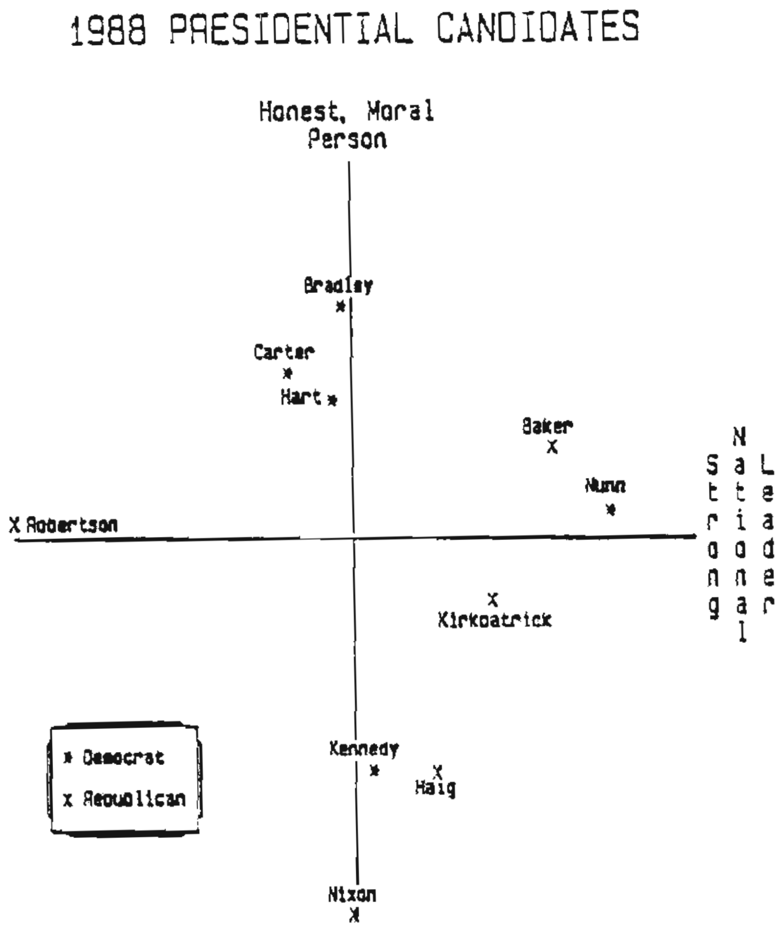
Figure 4

1988 PRESIDENTIAL CANDIDATES



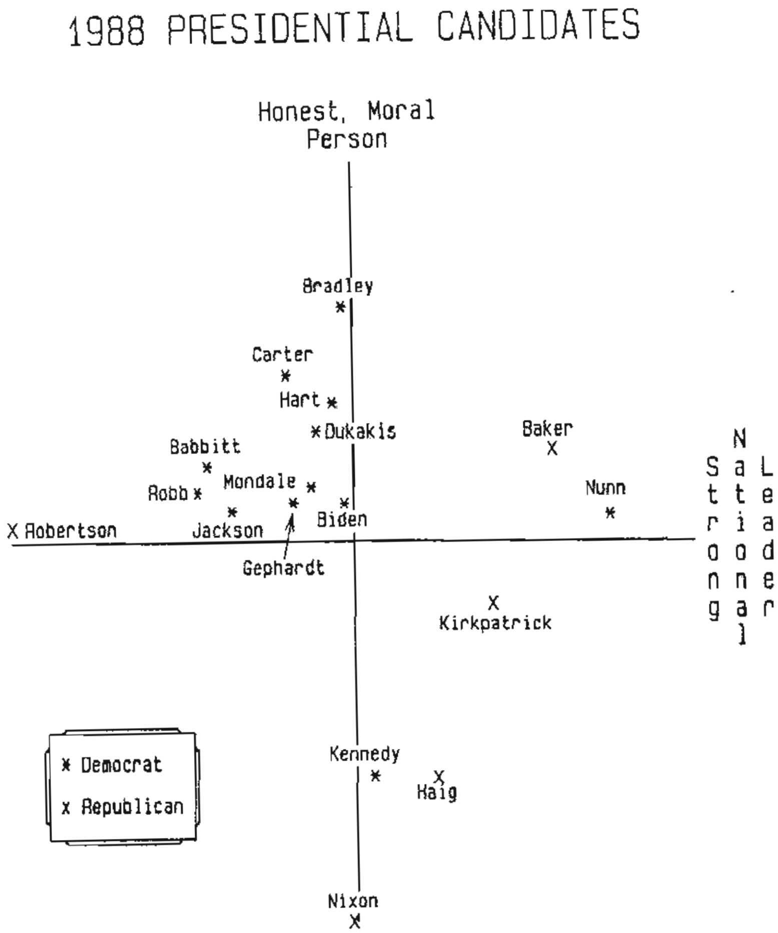
These candidates define the extremes of the moral aspects of character. Which candidates are perceived as strong national leaders? As shown in Figure 5 Pat Robertson is not. While he may be seen as a strong leader in other areas, in terms of executing the office of the Presidency, negotiation, and having a strong understanding of national and international issues, he is the antithesis of a "Strong National Leader." Three candidates scored very high on this dimension: Howard Baker, Sam Nunn, and Jeane Kirkpatrick.

Figure 5



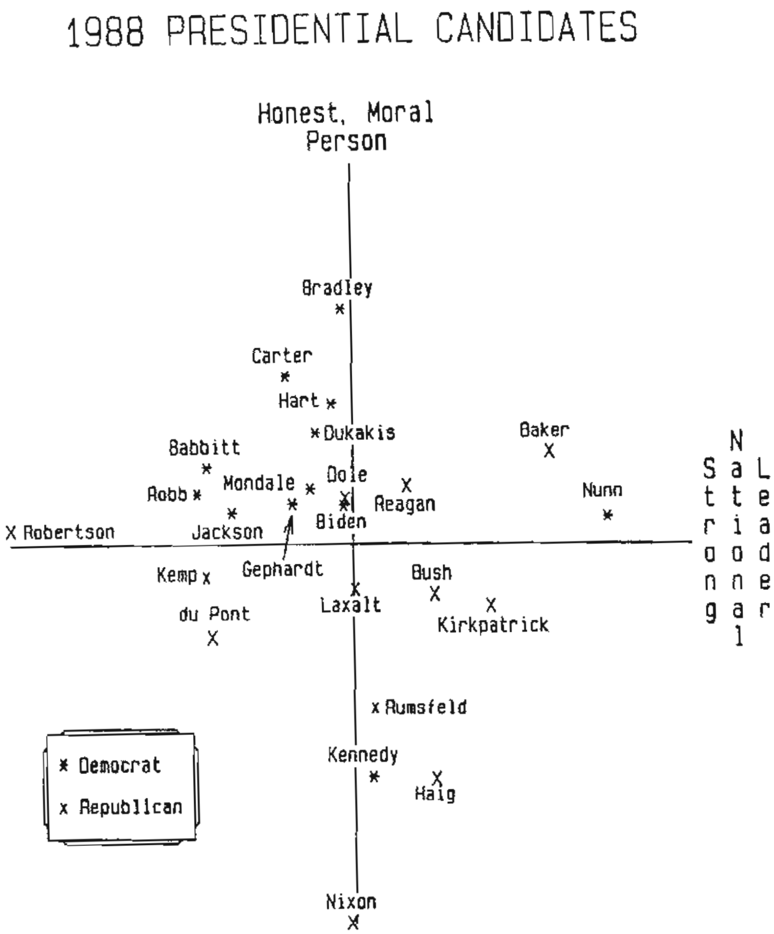
The Democrats as a group are seen as less knowledgeable and less strong in the area of "Strong National Leadership" with the exception of Sam Nunn. As a group they are seen as stronger on the "High Moral" scale (Figure 6).

Figure 6



When we add other Republicans, they cluster a little more to the right and somewhat lower on the "Honest, Moral Person" axis (Figure 7).

Figure 7



One of the advantages of dealing with a single product space is that once the audience understands that space, you can talk about each particular brand or product in that space and the particular perceptions of individual subgroups. You can look at the perceptions, for instance, of a set of candidates by respondents who classify themselves Democrats or Republicans. By dealing with a sample space, you can talk about all the products and their relationship to each other as perceived by a variety of different groups.

Rule #5 - Keep The Focus On The Screen

I have often been amazed by the impact of a perceptual map on an audience at the end of a meeting. Some of the most intelligent discussions about marketing and advertising strategy have taken place in the context of a single map. That kind of discussion cannot be generated if the audience has seen two maps. Someone is likely to say: "Let me see the contrast between the first and fourth dimensions," and you have to flip back through a pile of charts to clarify the question.

How hard is it to utilize this analytic technique? Having spent the whole day on the details and its complexities, our technical wits may have been scared out of us. The dangers of doing a perceptual mapping study are much more attributable to using bad fundamentals of marketing research, rather than a slightly inappropriate method of scaling or rotation. The mapping studies that work poorly are not done by good, solid practitioners of marketing research. Any solid practitioner of marketing research should be able, with today's technology, to take advantage of the powerful analytic and computational tools. In other words, we've built up an elitism about some of these techniques that is undeserved.

CONJOINT ANALYSIS: HOW WE GOT HERE AND WHERE WE ARE

by
Joel Huber
Duke University
Durham, NC

Conjoint analysis has had a profound effect on the conduct of research in many facets of business, particularly in the areas of product positioning and new product development. It is a field approaching the maturity stage of its life cycle. However, with the coming of inexpensive, user-friendly programs for conjoint, we can expect its use to increase substantially. Indeed, we will soon see the day when virtually all market research firms will offer conjoint studies as part of their standard repertoire. Managers will use conjoint not just for special projects, but as an indispensable tool enabling them to test the impact of proposed actions on the market. Conjoint is becoming less elite, its secrets no longer the property of a few, but available in its simpler versions to all.

Today I would like to present my personal perspective on the history of conjoint analysis. The field is shaped by two fundamentally conflicting forces. First, there are the idealistic psychometric forces that started the field. Opposing these, while at the same time arising from them, are the pragmatic forces, practitioners who have determined the way conjoint is used. The tension between these forces has shaped the growth of the field and will continue to guide its future development.

The Psychometric Tradition

The term "conjoint" has itself contributed to the mystery of the field. The term arose out of an attempt to apply extensive measurement to preference judgments. Extensive measurement refers to a method to build a scale by comparing relative lengths (extensions) of objects. For example, by comparing the lengths of different rods put end to end, one can form a scale on which it is appropriate to perform such operations as addition and subtraction. While such interval-level scales are relatively easy to generate from physical quanta such as weight, size and time, they have been notoriously difficult in the case of human preferences.

The difficulty arises because we know what it means to say that we like potatoes better than rutabagas, but generally not what it means to say that our liking for potatoes over rutabagas is greater than our liking for artichokes over eggplant. This indeterminacy poses a problem to psychometricians who want the same solid base for measuring the psyche as physicists had for measuring weight. Without interval scales of preferences, it is difficult to specify what it means to have an additive model of preference.

The psychometricians reasoned that while our ordinary language pronouncements of preferences do not directly produce interval scales, certain kinds of preference judgments had to be based on utility values that do. One set of preference judgments that requires metric underpinnings refers to compound or conjoint objects.

Consider the statement that one prefers a $10,000 convertible to an $8,000 sedan. This statement implies that the benefit of a convertible over a sedan is greater than $2,000. Psychometricians were able to show that by putting together a number of such preference statements, it is possible to derive intervally scaled additive partworth utilities that could underlie these preferences. Further, they specified a number of tests to determine if such an interval scale is justified, given the preference orderings.

Conjoint measurement provided a theory for creating a measurement scale from judgments on compound or conjoint objects. It generated a great deal of excitement when first proposed. Conjoint was a "psychometric conjurer's stone"- a way to transform the dross of ordinal preferences into the gold of interval scales. At last the measurement of preferences might be put on par with measurement in the exact sciences.

The early contributions focused on finding sets of elegant axioms and/or conditions required to uncover the latent interval partworths. Some of these conditions, such as independence, are well known, while others, such as double cancellation, are less well known. The axiomatizations are best summarized in the classic Foundations of Measurement, Volume 1 (1971) by Krantz, Luce, Suppes, and Tversky. In the preface to that volume, reference is made to Volume 2 on applications. It is ironic and significant that that volume has not yet been published.

What happened? As soon as the psychometricians applied their models to human behavior, they found that the axioms were consistently violated. It was very similar to what is now occurring with respect to the von Neumann-Morgenstern axiomatization of choice under uncertainty (Thaler, 1985). Virtually all the axioms were violated in relatively minor but systematic ways. Initially, it appeared that random error could account for the intransitivities and lack of additivity found. However, as more elegant and precise tests were devised, this escape was also blocked (e.g., see Falmange, 1976).

In hindsight, it is not surprising that if people cannot give consistent interval partworth values directly, then such metric rigor is unlikely to be hidden beneath more complex judgments on conjunctive stimuli. There is no intervally scaled ruler hidden in the brain that can account for complex preference judgments.

Still, the psychometricians provided a clear and coherent tradition, aspects of which are still important today. That tradition includes the following components: First, the belief that individual preferences can be expressed in numerical terms that lead to behavior. Second, the focus on comparisons among conjunctive stimuli, defined on multiple attributes, so that the response requires trading off high levels on one attribute with low levels of others. Third, the tradition of using factorial designs in which the attributes to be tested are statistically independent of one another. Fourth, the emphasis on testing the assumptions, such as additivity, as a prior condition to estimating the partworth utilities. Finally, the orientation to ordinal responses from subjects as the primitive behavior being modeled, rather than direct magnitude or interval scales.

From Psychometric Swords to Market Researcher's Plowshares

The psychometric tradition is rigorous and idealistic, whereas its adoption by the market research community has been approximate and pragmatic. The market research community began with the same rigorous models, but soon found that the partworth utilities were managerially very useful despite the fact that the tests did not work. In effect, the operation failed, but the patient thrived. Useful aspects of the original conjoint measurement framework were adapted and less useful ones were dropped.

Rich Johnson's succession of conjoint models is perhaps most illustrative of the changes that occurred. Rich was trained as a psychometrician, and his original trade-off analysis used 3-by-3 trade-off matrices, in which respondents were to rank order alternatives defined on various levels of the two attributes (Johnson, 1974). Then, by computerizing the approach, he was able to avoid certain redundant questions and speed up the task. However, the price of this additional speed was a lessened ability to make consistency tests at the individual level. His next step expanded the task from one of categorical preferences to graded-pair comparisons. This permitted more information to be collected from respondents with very little additional cost in time or effort. Finally, he used the personal computer to merge direct attribute judgments with paired comparisons and guide the selection of "optimal" pairs during the conjoint task. All these changes helped to obtain information from respondents more efficiently and to formulate a better predictive model of their preferences. Still, these changes represent a substantial departure from the psychometric tradition.

Much of the ambivalence between idealism and practice in the marketing research community is found in Green and Srinivasan's (1978) classic review article on conjoint. In that article they differentiate conjoint analysis from the older conjoint measurement in order to make appropriate separation between the two fields. A dualism is evident in their discussion of the various ways to perform conjoint analysis, sometimes focusing on what is theoretically justified, while at other times succumbing to practical reality.

What did the marketing research community take from the psychometricians and what did they change? Generally, the trends are evident in Cattin and Wittink's (1987) review of practices in conjoint. First, the field continues to consider behavior as captured by partworth utilities and simple additive models. Second, in keeping with the psychometric tradition, they use compound stimuli which force individuals to trade off conflicting attribute levels. Third, they still rely on orthogonal arrays, although highly fractionated designs have replaced the original full factorials. The first three components of the psychometric tradition have been passed down relatively unchanged. The last two, the structural tests and the nonmetric orientation, have rapidly eroded.

Consider, first, the tests of the structural composition rules (such as additivity) that were the major focus of the axiomatic systems. These tests are now virtually ignored, or worse, assumed away. Consider, for example, the common use of fractional main-effects design. While these offer far greater efficiency and permit main-effects estimates of many more attributes, they assume that interactions are zero. If there are interactions, the preference function will be biased or wrong. Further, because no test is possible, the analyst will never know that the results are biased.

The other major deviation from the psychometric tradition has been a move from a nonmetric to a metric orientation. This has occurred both in the kinds of data collected from respondents and routines used to analyze it. The original reason to use rank-order inputs over quasi-metric ones stemmed from a legitimate uncertainty about what respondents meant in responding to, for instance, a ten-point strength-of-preference scale. A number of nonmetric analysis packages, such as Kruskal's (1965) MONANOVA and Srinivasan's and Shocker's (1973) LINMAP permitted relatively easy analysis of input data about which only ordinal properties could be assumed. The shift from ordinal to metric inputs has been largely pragmatic. For example, putting 25 profiles onto a ten-category sort board is both easier for subjects and provides more reliable inputs than an exhaustive rank-order task. Using a rating scale permits one to generate predicted choices with equivalent reliability but fewer judgments.

Metric methods have also become more popular as methods of analysis. This shift is in part due to the ease of use of ordinary least squares over nonmetric routines. But it also stems from a realization of the value of the weak but errorful metric information in a rating or a sort-board task. Of course, such data can be analyzed by nonmetric procedures, such as monotone regression or LINMAP. The nonmetric procedures find a monotone transformation of the dependent variable that best fits the model. With such routines this transformation is very linear, indicating that the transformation provides little additional information. More significantly, the monotone transformation generally degrades the predictive ability of the model (e.g., see Huber, 1975). Quasi-metric data has some interval properties that nonmetric routines treat as noise, but metric routines are able to use. Nonmetric routines may be losing popularity simply because they do not appear to help predictions. However, this is simply a pragmatic criterion; we lack a good theoretical reason why metric routines should work better. Indeed, most theoretical considerations lead to the championing of a nonmetric orientation.

To summarize, the marketing research industry has taken some of the external trappings of conjoint measurement, but has generally deleted or modified its elegant inner workings - those dealing with hypothesis testing on strictly ordinal inputs. While these modifications may not, in a strictly predictive sense, matter, it is appropriate to inquire as to the kind of offspring that has emerged from this union of pragmatic and idealistic parents.

<u>What Really Goes on During a Conjoint Exercise?</u>

If we are not tapping a latent interval utility scale during conjoint, what are we doing? The answer depends on whether the stimuli are unitary or decomposable. Unitary stimuli are those which respondents cannot easily break into component parts, such as foods, scents, or esthetic objects. With these kinds of stimuli, response to factorially designed stimuli is holistic and generally frustrates attempts to build simple models of that response. Main-effects designs are at a particular disadvantage in that the assumption of additivity is usually violated. Specialized designs are appropriate if the source of the interactions can be localized within a few variables, and if there is not too much heterogeneity across subjects, conditions that are sadly not often satisfied.

Because of the design problems with unitary stimuli, most conjoint analysis uses readily decomposable attributes and displays attributes in ways that make it easy for respondents to separate them. Respondents are given repeated questions with predictably arrayed attributes, and anyone who has watched a conjoint exercise knows the result. Respondents simplify the task by focusing on a few attributes. Each profile is evaluated by scanning these attributes and adjusting the valuation of the alternative accordingly. This process results in a very good fit of the additive model at the individual level. Typically, a small number of attributes are strongly significant and the rest are nonsignificant. Interactions are very rare - they require extra processing. Different tasks produce slightly different patterns of responses. For example, paired comparisons may produce more significant attributes. However, the general pattern of a strongly simplified evaluation strategy that has accurately captured an additive model emerges regardless of the particular task.

Evidence for such a simplification process appears in an anomalous result which I have found in my work, and I suspect many of you have as well. Following a conjoint exercise we often include a choice task, either using actual brands or profiles where the attributes have been scrambled to reduce the likelihood that the conjoint choice process will be trivially repeated. In these studies it is possible to note the correspondence between the internal fit of the conjoint and its accuracy in predicting the holdout choices. If the conjoint task is a correct representation of the holdout choices, then the better the fit of the conjoint and the more it satisfies the axioms, the better it should predict the holdouts. In other words, respondents who are well modeled by the conjoint should be ones we can best predict. However, my experience has not been consistent with this expectation. Correlations between conjoint fit and predictive accuracy are very low and often negative, particularly if one screens out the totally random subjects. How could this be?

The simplest account focuses on differences in the conscientiousness of respondents. The conscientious ones try to consider as many attributes as possible in their conjoint task. Being mortal, they make mistakes, resulting in preference reversals and greater error levels. When they make the holdout choices, they conscientiously try to be consistent with their earlier judgments, resulting in greater predictive accuracy. For the less conscientious respondents the reverse holds. They simplify the conjoint task greatly in order to get through it, focusing on one or two attributes. This simplification permits remarkable fits to be the additive conjoint model. In the holdout choice task with distorted or scrambled attributes, using the same simplified decision rule is not easy. These less conscientious respondents shift strategies, basing their choices on different attributes than in the conjoint, resulting in a poor correspondence between the two.

The point is not to resolve this anomaly, but to raise an issue about conjoint. The professional success of conjoint practitioners attests to how well it works. It produces intuitively pleasing results that managers find very useful, although we do not have a clear account of why it works. Indeed, its problems could lead many to discard it. After all, it has shed its theoretical roots and appears only to capture a simplified and truncated version of choice behavior. The next section considers why conjoint works, and this leads naturally into ways in which it might be improved.

Why Conjoint Works

1) Conjoint requires tradeoffs that are similar to those in the market. A conjoint task is valuable because it forces the respondent to evaluate conflicting attributes, as between the type of car and its price. People typically try to avoid

making such judgments by searching for unambiguous solutions involving dominance or following a relatively clear heuristic. However, the marketplace also requires such judgments and people make them when they must in the marketplace or the conjoint task.

The conjoint task, in which alternatives are compared on a number of dimensions, can be usefully contrasted with a direct elicitation approach, in which attribute utilities are directly assessed. There are two problems with direct elicitation that are ameliorated with conjoint. First, it is difficult with the direct elicitation approach to keep a respondent from seeing everything as important. Certainly, $2,000 is very important in selecting a car, but it may not be more important than the difference between a convertible and a sedan. Second, direct elicitation does not directly relate to a choice in the marketplace, but is a summary measure of those behavioral decisions. In contrast, the conjoint task is more directly analogous to market choice.

2) The simplification in conjoint may mirror that in the market. The simplification found in conjoint to a small number of attributes is only misleading if there is a very different kind of simplification in the marketplace. There is evidence that the decisions in the market are based on remarkably few dimensions (Olshavsky and Granbois, 1979). If so, then conjoint may indicate those few attributes on which the consumer bases his or her decisions.

Further, to the extent that conjoint is used to predict aggregate shares, it does not matter that an individual's selection of attributes is unstable over time. As long as conjoint captures an unbiased selection of attributes at the time, the aggregate market shares will also be unbiased. The criteria for conjoint to work at the aggregate level are considerably less stringent than for individuals.

To summarize, conjoint works by forcing respondents to make trade-offs among attributes. They then simplify the task by selecting a small number of attributes on which to base their judgments. To the extent that this pattern of simplification is mirrored in the marketplace, then conjoint market shares will predict quite well.

3) Conjoint profiles are orthogonal. The use of orthogonal arrays is an aspect of the original psychometric formulation that has resisted modification by the marketing research community. In particular, main-effects fractional factorials have been heavily used because they permit more attributes and levels. In the case of decomposable stimuli, the simplification by respondents typically assures that interactions will not be present.

The orthogonal nature of these designs is important in a way not generally appreciated. An orthogonal design is simply one in which the levels of different attributes across profiles are uncorrelated. Such designs assure that an estimate of one attribute is unaffected by the estimate of other attributes. It might appear that we could suffer moderate levels of multicollinearity without much harm. Most econometric models seem to thrive with much higher levels of multicollinearity. However, the fact that respondents regularly simplify the conjoint task leads to substantial difficulties if any attributes in the design are correlated. Let me illustrate with an example.

I was involved in a conjoint study dealing with snowmobiles. We were concerned with the impact of engine size on the acceptability of the snowmobiles. So that the profiles would be realistic, we increased the price by an appropriate amount for each of the engine sizes. Price was positively correlated with engine size in the design. This was not expected to be a problem, except that multicollinearity might render the estimates marginally less efficient. However, we found that for many respondents the coefficient of price had the wrong sign (high price is preferred) and for others the coefficient of engine size had the wrong sign (small size is preferred). We believe that respondents, in simplifying, had tended to focus on one of these two variables. For example, those who focused on engine size gave higher evaluations of profiles with larger sizes, but these, by our design, had higher prices. Thus, high price appeared to be desired by these subjects.

This account points to an advantage inherent in orthogonal arrays. For orthogonal arrays, the main-effect estimate for each attribute is independent of the others, whereas in the correlated case, this independence does not hold. If attributes are correlated, misspecification results in biased estimates. The particular misspecification that so often occurs in conjoint is simplification, where a number of attributes are effectively ignored. With orthogonal arrays, the estimated coefficients for the attributes remain unbiased. In the correlated case, misspecification results in distortions in the coefficients for both the attributes focused upon and those ignored. Thus, orthogonal arrays play an important role of increasing the robustness of conjoint by making it less likely that coefficients have counter-intuitive signs. This robustness contributes to much of the managerial satisfaction with conjoint.

4) <u>Conjoint Simulators Account for Heterogeneous Tastes in a Market</u>. A final reason conjoint works relates to the way it is used. Typically, partworth functions are estimated at an individual level, then these are aggregated to produce estimates of market share under various conditions or scenarios. These simulations implicitly reject the notion that one homogeneous customer can account for choices in the marketplace. Instead, one is forced to deal with each customer having an idiosyncratic preference function, or at least with an explicit clustering in which strongly differing tastes are represented in different clusters. This practice of preserving the heterogeneity of individuals in simulators facilitates the representation of two important properties of markets that are difficult to achieve with other market research techniques. These are the properties of differential substitution and dominance.

Differential substitution refers to the notion that a new competitor in a market tends to take share differentially from those brands with whom it is most similar. For example, New Coke took most share from Classic Coke and Pepsi, and had relatively little impact on the lemon-lime soft drink category. Dominance refers to the idea that a brand that is equal on most attributes but slightly worse than its competitor on others gets very low share.

Managers understand these two phenomena and expect simulations to reflect them in positioning and new-product studies. Unfortunately, most models of market structure can account for differential substitution and dominance only in a very awkward fashion. By contrast, both phenomena arise naturally out of a conjoint simulator. For example, differential substitution occurs because individuals who like Pepsi tend to like Coke. Generally, changes in any brand will have a greater impact on similar brands than dissimilar ones. Dominance is represented in a conjoint simulator since a brand that is dominated consistently loses out to that competitor and achieves almost no share.

To summarize, conjoint works because it is derived from a task that forces respondents to trade off attributes in ways that may parallel actual buying behavior. The orthogonal design provides not only efficiency, but a strong degree of robustness against misspecification. Finally, the preservation of the utility function at the level of the individual or segment permits us to simulate a market that behaves in ways we expect.

The Future of Conjoint

There are three areas in which substantial changes are anticipated. The first area of change involves conjoint theory, the way we think about and organize the field. The second involves the task. The final area involves the ways conjoint is used.

1. Conjoint Theory: From Estimation of Utilities to Emulation of Behavior In terms of theory, the psychometric framework has one fatal flaw - it assumes that utilities exist that account for preferences. Reality, unfortunately, is far more complex. Preferences between profiles are better described as being constructed, using various heuristics from the information at hand. The additive models in conjoint may reflect this process and at times correspond to it quite well, but conjoint certainly cannot reveal an interval scale in the brain. Instead of thinking that conjoint estimates latent utilities, it is more appropriate to consider that it emulates choice behavior.

There is a significant loss associated with giving up the idea that conjoint reveals a latent preference scale. The existence of a scale means that it is possible to formulate optimal experimental designs that result in the most efficient estimates of that preference scale. If, instead, conjoint is viewed as a paramorphic emulation of behavior, then it is no longer clear what makes a good design.

There are some advantages with viewing conjoint as an emulation of choice behavior. First, we are no longer permitted to beg the question about the applicability of conjoint to the marketplace - something we can do if the same utility scale is presumed to underlie both choice and conjoint. Instead, we are forced to ask whether the conjoint task corresponds to the choice in the marketplace. For example, it is relevant to assess the number of dimensions that are actually used in the market, then choose a conjoint task that results in similar depth of processing. Second, in the behavioral perspective, one is freed from a rigid adherence to a particular question form to capture appropriately the choice process.

In contrast to utility, behavior can be captured in a number of paramorphic ways. Indeed, to get at some behavior it may be better to use different kinds of questions, such as combining direct elicitation and paired comparisons, rather than focusing on a particular question type. Switching to a different kind of question may discourage respondents from getting into a response pattern. Further, the differences in responses across question types will reveal the stability of the choice behavior.

2. <u>Conjoint Task: From Monolithic and Rigid to Multifaceted and Adaptive</u> As our way of thinking about conjoint changes, so will the task we ask of subjects. Two important changes will occur, both adding new kinds of questions to the traditional conjoint task. The first involves the ability of routines to adapt to the idiosyncratic behavior of respondents, while the second adds a relatively realistic choice task at the end of the conjoint task to better assess the correspondence between conjoint and market choice.

We have already begun to see ways in which conjoint can profitably adapt to the needs of individual respondents. The Sawtooth Software ACA System uses "priors" to construct pairs so that paired comparisons are as closely balanced as possible. While this reduces the strict statistical efficiency of the design, it makes the questions more challenging and increases the correspondence between conjoint and subsequent choice (Huber and Hansen, 1986). The flexibility of the personal computer in administering conjoint will certainly lead to other adaptive mechanisms. Two such applications are particularly exciting - hierarchical conjoint and interaction testing.

Hierarchical conjoint permits the modeling of rich decision making despite simplification of the conjoint task by respondents. If respondents can only cope with two or three attributes at a time, the routine determines the partworth functions for these most important attributes, then fixes their levels. Subsequent test profiles only differ on the remaining attributes. For example, the values of location and price might be estimated first in an apartment study, followed by furniture style and room layout. Under standard conjoint, these less important attributes might not be revealed, whereas in hierarchical conjoint both their position in the hierarchy and relative importance could be assessed.

A second adaptive mechanism concerns the search for interactions. A promising technique might work as follows: First, a main-effects design would make rough estimates assuming no interactions. The residuals from these initial judgments would be tested for weak (and confounded) evidence of interactions. Then these potential interactions would be tested through specially designed questions. Such a method would avoid the current problem of having to assume that interactions are zero, and could be very helpful in studies where different interaction patterns are expected across respondents.

Perhaps the greatest prospect for improving conjoint involves the inclusion of a relatively realistic choice task at the end of the exercise. These choice tasks sometimes take the form of asking respondents to choose brands from a simulated store or having them evaluate actual products. A much less costly, if somewhat less realistic option, is to add choice questions at the end of the conjoint exercise. These might take the form of choices among alternatives defined on different attributes than those displayed in the original conjoint. This is easy to do within the Sawtooth Software ACA System by adding Ci2 System questions after the ACA System section. Further, new developments in personal computers will permit potential choice objects to be displayed in color video, thereby increasing the realism of the holdout choices even more.

The value of having a holdout choice task is twofold. First, in the field this holdout task is useful in identifying respondents whose conjoint responses are unlikely to correspond to their behavior. These respondents can then be given less weight in the simulation. Second, it permits an immediate assessment of the relevance of conjoint to choice. Where a conjoint model appears not to correspond to choice, it can be changed or improved. This permits the testing of different forms of conjoint and leads to continuous improvement in its predictive validity.

3. <u>Conjoint Simulators: From Complex and Opaque to User-Friendly and Understandable</u> The third area in which we can expect conjoint to progress involves the way data are used in choice simulators. Elaborate choice simulators currently exist, permitting the analyst to ask virtually any question of the data. The positioning of a product can be optimized with respect to sales or profitability. Alternatively, one can assess the impact of changes on positioning or competition on the behavior of various segments. Unfortunately, these simulators are not particularly user-friendly. With time they will become easier to use and their use by managers should increase.

Even if made more user-friendly, there is still a problem. While simulators permit the manager to cope with heterogeneous tastes in the market, they remain a black box. The only way for a manager to understand a simulated market is by experimenting with a large number of runs. These simulation runs give managers a feel for the behavior of the market in the face of different positionings or competitive offerings. Developing this understanding is hard, relatively unstructured work. We need to develop ways to permit managers to understand more directly the behavior of the market being simulated. Preference spaces may provide part of the answer, but it is very difficult to represent both respondent heterogeneity and central tendency on one map. Working with a small number of segments also helps. We would like to know how to specify a small number of segments, such that their aggregate behavior closely approximates the entire market. Defining such segments remains an unsolved question that evades simple solutions.

Just as the computer continues to make the conjoint task itself more appealing to respondents, it will also increase the ease by which the output of conjoint can be applied by managers. Once again, we may have been saved by the computer. We have come to realize that choice behavior cannot be captured by a simple scale of utilities. As we come to accept a behavioral base to conjoint, we can never return to the elegance and simple unity that characterizes the psychometric framework. However, with the computer we have a tool that may be powerful enough to mirror the complexity of behavior and display it in a manageable and understandable way.

REFERENCES

Cattin, Philippe and Dick R. Wittink (1987) "Commercial Uses of Conjoint Analysis: An Update" Working Paper, Graduate School of Business, Cornell University, Ithaca NY.

Falmange, Jean-Claude (1976) "Random Conjoint Measurement and Loudness Summation" Psychological Review, 83, 65-97.

Green, Paul E. and V. Srinivasan (1978), "Conjoint Analysis in Consumer Behavior: Issues and Outlook" Journal of Consumer Research, 5, 103-23.

Huber, Joel (1975) "Predicting Preferences on Experimental Bundles of Attributes: A Comparison of Models" Journal of Marketing Research, 12, 290-7.

Huber, Joel and David Hansen (1987) "Testing the Impact of Dimensional Complexity and Affective Differences in Adaptive Conjoint Analysis" Advances in Consumer Research, 13, in press.

Johnson, Richard M. (1974) "Trade-Off Analysis of Consumer Values" Journal of Marketing Research, 11, 121-217.

Krantz, D.H., R.D. Luce, P. Suppes and A. Tversky (1971) Foundations of Measurements, Vol 1, New York: Academic Press.

Kruskal, Joseph B. (1965), "Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data" Journal of the Royal Statistical Society, Series B, 27, 251-63.

Olshavsky, Richard W. and Donald H. Granbois (1979), "Consumer Decision Making--Fact or Fiction" Journal of Consumer Research, 6, 93-100.

Srinivasan, V. and Allan D. Shocker (1973) "Linear Programming Techniques for Multi-dimensional Analysis of Preferences" Psychometrica, 38, 337-69.

Thaler, Richard (1985), "Mental Accounting and Human Choice" Marketing Science, 4, 199-214.

ADAPTIVE CONJOINT ANALYSIS

by
Richard M. Johnson
Sawtooth Software, Inc.
Ketchum, ID

Joel Huber has provided an impressive discussion of conjoint analysis. My focus will be narrower, and I will describe a particular approach to conjoint analysis. I will also stay at a non-technical level; more detail is available in a paper, "Adaptive Conjoint Analysis," from Sawtooth Software.

Conjoint analysis assumes that products are decomposable into separate attributes. In a conjoint study we show respondents many hypothetical product concepts that differ systematically in their attributes, and ask respondents for overall reactions to each concept. From their responses, and our knowledge of the attributes composing each concept, we try to infer the values respondents place on the separate attributes. Conjoint analysis is, therefore, only appropriate for product categories where the "value" of a product can be approximated by adding up the "values" of its various parts.

If we find that product preferences can be accounted for in terms of the known values of their attributes, then we might expect to predict preferences for new products consisting of different combinations of the same attributes. Surprisingly, this idea has turned out to work rather well, and conjoint analysis has been of value in a wide range of product categories.

As a first assumption, then, we think of products as "bundles of attributes," each with specified levels. For example, if we were interested in laptop computers we might consider the attributes shown in Figure 1.

Figure 1

ATTRIBUTES FOR LAPTOP PCs

SCREEN LEGIBILITY
- Very Easy to Read
- Easy to Read
- Hard to Read

WEIGHT
- 9 pounds
- 12 pounds
- 15 pounds

SPEED
- Like IBM PC
- Twice as fast as IBM PC
- Four times as fast as IBM PC

BATTERY LIFE
- 12 Hours
- 6 Hours
- No Battery; Must Plug In

PRICE
- $ 900
- 1,900
- 2,900
- 3,900

The second assumption is that each attribute level has a particular value for a respondent, which affects how much he "likes" a product. We call these values "utilities" or "partworths."

Figure 2 presents values for a hypothetical respondent. The least-liked level of each attribute has a value of zero.

Figure 2

A HYPOTHETICAL SET OF UTILITIES

Attribute Level	Utility
Very Easy to Read	8
Easy to Read	7
Hard to Read	0
Weighs 9 pounds	3
Weighs 12 pounds	2
Weighs 15 pounds	0
4 x Speed of IBM PC	7
2 x Speed of IBM PC	2
Speed of IBM PC	0
12 Hour Battery	3
6 Hour Battery	2
No Battery; Plug In	0
Costs $ 900	10
Costs 1,900	8
Costs 2,900	5
Costs 3,900	0

The important information is contained in the differences between these values. Consider screen legibility. For this respondent, "Very easy to read" is worth 8 points, and "Easy to read" is worth 7 points. This respondent would experience one additional unit of utility in improving from merely "Easy" to "Very easy" to read. He would experience the same difference in utility if battery life were improved from 6 hours (worth 2 points) to 12 hours (worth 3 points), also a difference of one point.

Since the least-liked level has a utility of zero for each attribute, the highest utility for each attribute is equal to that attribute's range from the best to worst. That range (the difference in utility between best and worst levels) represents the maximum impact in utility that the attribute can contribute to a product. This is sometimes taken as an expression of the "importance" of the attribute. Price is particularly important for this respondent, with a range of 10. Battery life and weight are least important, both with ranges of only 3.

Our third assumption is that the utility of a product is just the sum of the utilities of its attributes. In Figure 3 the utilities for two products are obtained by adding up the utilities of their attribute levels. Based on these utility values, this respondent should prefer Product B to Product A.

Figure 3

Product A		Product B	
Hard to Read	0	Very Easy to Read	8
9 pounds	3	12 pounds	2
2 x PC Speed	2	4 x PC Speed	7
12 Hour Battery	3	6 Hour Battery	2
Costs $900	10	Costs $2,900	5
	--		--
Total Utility	18	Total Utility	24

The fourth assumption is that we can reverse the previous process. Instead of adding the utilities and predicting which product a respondent ought to prefer, we can show him a carefully designed group of product concepts, inquire about his preferences, and then infer his underlying utility values. Three methods have been most popular.

First, most conjoint studies have used the "full-profile" method. With this method, a set of concepts or "profiles" is composed in which every concept is fully specified (that is, its level is specified for every attribute). The concepts are carefully chosen so that combinations of attribute levels are presented in a balanced way. If the concepts are well chosen it is possible to infer the respondent's utility values for attribute levels from his overall preferences for the concepts.

With the full-profile method we may ask the respondent to rank the concepts from best-liked to least-liked. In the early days of conjoint analysis, many studies were done that way. A nonmetric method would most likely be used to estimate utilities from ordinal data. However, there is more recent evidence that ratings on a numeric scale, perhaps ranging from "like very much" to "don't like at all," can provide equivalent data with less effort for the respondent. Rating scale data are usually analyzed by regression or other metric methods.

An advantage of the full-profile method is that the task is very natural. You're showing the respondent concepts that are like real products, in the sense that they're described on all the relevant attributes. The respondent doesn't have to keep an "all else being equal" mental set, as he would if the concepts were described on only some of the attributes.

However, the full-profile method does present some problems. First, respondents tend to simplify the task. Suppose we had 12 attributes instead of five. Each concept will have a dozen aspects, and there are likely to be as many as three dozen concepts to consider. Most respondents are not capable of reading that much material conscientiously. A common way of coping with such a difficult task is to choose a few attributes as most important, and to attend only to those.

Another problem with full-profile analysis is that it's hard to accommodate a design large enough to be practical in the real world. It's fine with four or five attributes, and may be usable with as many as eight. But it doesn't work very well with a dozen. A study of as many as 20 attributes would be completely absurd.

The first conjoint study I was involved in occurred before I had even heard the word "conjoint." A researcher working on the design of a new product commissioned a concept test involving 30 attributes. The project took a month. Meanwhile, his product manager had found it necessary to make the decisions that were being researched without waiting for the results of the test.

This happened again and again. Finally the researcher asked:

> "Please, find some way to do a study once, and file the results in a computer, so that when my product manager says he wants a concept test, I can just do a run on the computer and give him the answer right away!"

That request led to the method of trade-off matrices (which since that time has become nearly obsolete). That method can deal with many attributes. A trade-off matrix asks a respondent to consider a pair of attributes. It displays all combinations of levels for those attributes, asking the respondent to indicate his rank order of preference for the combinations. The respondent is asked to write a "1" in the cell corresponding to his first choice, a "2" for his second choice, etc. A trade-off matrix is shown in Figure 4.

Figure 4

A TRADE-OFF MATRIX

	BATTERY LIFE		
SPEED:	12 Hrs	6 Hrs	None
4 times IBM PC			
2 times IBM PC			
Equal to IBM PC			

One advantage of trade-off matrices is their ability to handle large numbers of attributes. It is not necessary to compare each attribute to every other one. The number of trade-off matrices presented has normally been between 1.5 and 2 times the number of attributes. A respondent can fill in the boxes quickly, and even 30 attributes does not present a hopelessly lengthy task. Another advantage is that, unlike the full-profile method, the trade-off method forces the respondent to pay attention to every attribute.

But there are also some serious problems. First, the task is extraordinarily artificial and many respondents can't even understand what you want them to do. Second, the task can become tedious with a large number of attributes; respondents get blurry-eyed after too many matrices. After fatigue has set in they tend to simplify the task by mechanically filling in numbers across rows or down columns. Others place a "1" in the middle of the box, then fill in around the outside with 2, 3, 4, etc.

Yet another problem is due to the fact that only two attributes are shown at a time, and the respondent must remember that "all else is equal." Failure to do this can lead to problems such as having respondents prefer a higher price to a lower price because they believe better products are likely to be higher priced.

In our Adaptive Conjoint Analysis (ACA) System, we've tried to retain some of the strengths of both the full-profile and trade-off matrix approaches. The ACA System proceeds as follows:

First, we screen all attributes and levels to discover some fundamentals about the respondent's preferences. This allows us to make preliminary estimates of his utilities, and also to eliminate attributes that are relatively unimportant to him and levels that would be unacceptable.

Next, we ask "trade-off" questions based on attributes important to the respondent. We use his answers to refine our estimates of his utilities, and also to choose further questions most likely to provide additional precision in our estimation of his utilities. This section of the interview is "custom-designed" for the respondent, and the questions asked depend upon our current estimates of his utilities, which depend in turn on his previous answers.

Finally, we ask "calibrating questions." These are like "full profile" questions, asking for his level of interest in several concepts specified on several attributes.

The "adaptive" nature of the interview lets us learn more about the respondent's utilities than we could otherwise. We can cover many attributes at many levels with a relatively small number of questions. Next, I'll describe a typical ACA System interview. It consists of several sections:

1) We show the levels of each attribute and ask, "Are any of these so totally unacceptable that you would not consider such a product under any circumstances?" The purpose of this section is to discard any levels which the respondent would reject so resoundingly that there is no point in asking further questions about them.

These questions can be dangerous because respondents are often too willing to discard levels. So few attribute levels may survive that it will be impossible to describe any real products using them. The researcher has the choice of including this section or not. We do not recommend including these questions unless the interview would otherwise be too long.

2) Next, we ask the respondent to rank the levels of each attribute for preference. The respondent chooses his most favorite level, next most favorite level, and so on.

Attributes which have a clear a priori order, such as "quality," "durability," and "effectiveness," need not be asked. If the researcher is confident that he knows how respondents would answer for an attribute, he may skip those questions.

3) Next we ask an "importance" question for each attribute.

"Importance" is a slippery idea. When you were deciding how to travel to this conference you probably chose an airline. Two attributes of airlines are fare and quality of food. The fare was probably more important in your decision than the food. That may have been because there were greater differences in the fares available than in the quality of the food.

If the only fares had been $100 and $101, fare would probably not have seemed very important. And if one meal in ten were known to cause ptomaine poisoning, then food would probably have seemed quite important.

How important something is depends on what's available. Our approach is to show the most preferred and least preferred levels for each attribute and ask, "If you couldn't have (blank) but had to settle for (blank) instead, how unhappy would you be?" The respondent answers on a numeric scale for each attribute.

This gives us an indication of the relative importance of each attribute, where we know which attribute levels are being contrasted in making those judgments. This allows us to focus on the most important attributes and also helps us formulate preliminary estimates of the respondent's utilities.

4) Next we ask a number of "trade-off" questions, an example of which is shown in Figure 5.

Figure 5

A TRADE-OFF QUESTION

Which would you prefer?

Four times Speed of IBM PC and Six Hour Battery	OR	Twice the speed of IBM PC and Twelve Hour Battery

Choose a number from the scale below.

Strongly Prefer Left 1	2	3	4	No Preference 5	6	7	8	Strongly Prefer 9 Right

The questionnaire author may choose how many trade-off questions are asked, and also how complex they are. The product concepts in the boxes can be composed of from two to five attributes. In telephone interviewing, it seems sensible to keep the questions as simple as possible, showing only two attributes at a time. In personal interviewing, we suggest starting with two attributes at a time, and then moving to three attributes after the respondent is comfortable with the task. It seems pointless to go beyond three attributes, since the increased complexity of the task seems to outweigh the value of the richer concepts.

Following each response the computer updates its estimates of the respondent's utilities, and uses this new information to choose the next question to ask. The questioning is ended by an automatic stopping rule unless the researcher has specified the number of trade-off questions that should be asked.

5) The final section presents "calibrating concepts." The computer automatically constructs several custom-designed concepts for the respondent, based on his previous responses. The concepts are designed to range from very undesirable to very desirable.

The researcher specifies the number of concepts to be asked and the number of attributes appearing in each. The respondent is asked for a likelihood-of-buying rating for each concept. This information is used to scale his utilities so that they contain information not only about his preferences, but also about his level of interest in the product category.

Details of the procedure for estimating utilities are given in the technical paper. We make preliminary estimates of the respondent's utilities before the trade-off section to serve as priors for a Bayesian updating scheme. At the conclusion of the trade-off section the utility estimates are true least-squares.

I have a personal bias in favor of nonmetric methods of estimating utilities. The ACA System is designed to permit those, but only after the interview is completed. Nonmetric estimates would take so long to update after each answer that the delays between questions would be too great. The ACA System produces a file containing a record of every question asked along with its answer. That information can be run through a nonmetric estimating routine to sharpen the utilities (or perhaps dull them; there is no convincing evidence that nonmetric procedures are better or worse than least squares).

To choose the next trade-off question, we use an algorithm that gives a relatively balanced design and also tries to ask difficult questions. Questioning is most efficient if the two concepts in a pair are chosen to be nearly equal in utility.

One advantage of the ACA System is that the researcher doesn't have to create the experimental design. He just needs to decide on his attributes and levels and to indicate whether he wishes to override any of the defaults built into the system. Another advantage is speed. The utilities are computed during the interview so you can begin your analysis as soon as the last interview is completed.

The adaptive nature of the process means that you can start with as many as 30 attributes, each with nine levels. The researcher indicates how much trimming should take place. In the extreme, you could ask the respondent to trade-off only the two or three attributes most critical to him.
Utility estimates are provided for attribute levels that are eliminated, although these are much less precise than for the levels studied in detail.

There is quite a bit of evidence that respondents remain more interested in the ACA System interview than they do in noninteractive tasks. Some have commented that the interview is a like a chess game. The computer makes a move, they give a response, and the computer is evidently taking that response into account when making its next move. The respondent can see that the machine is paying close attention, and asking more and more difficult questions.

The ACA System also provides a simulation capability. Simulations reverse the utility estimation process, predicting respondent preferences from their utilities. Preferences are predicted for hypothetical markets, consisting of several product concepts defined in terms of attribute levels. You get average shares of preference, standard errors, and average utilities. Simulations can be done with all respondents, or with subgroups defined in terms of demographic or other criteria.

ACA System simulations are always done at the level of the individual respondent, and then group results are obtained by cumulating shares of preference over individuals. There are several "choice models" available for estimating each respondent's shares of preference among the products in the simulation.

> The "first choice" model assumes the respondent will select the product having the highest utility for him. All his "share" goes to that product. It's a simple and popular choice model. In my experience, though, it tends to be biased, overpredicting for powerful concepts and underpredicting for weaker ones.
>
> The second model is a "share of preference" model. We use a logit transform of the utilities to estimate the relative liking for each product, numbers that are similar to probabilities. Those are divided by their sum to get shares of preference. Of course, they are not good estimates of "market shares," which would have to include effects of advertising, distribution, etc.
>
> The share of preference approach has a widely-recognized defect. If the researcher made the mistake of entering a product into the simulation twice, that product would receive two shares. Their total wouldn't be quite twice as great as it should be, but it would be too large. The same phenomenon extends to products that are similar to one another, but not identical. The more a product is similar to other products, the more its share of preference will be over-estimated.

Our third model makes a "correction for similarity." The correction is based on an elegant idea suggested by Dick Smallwood, also on this program. In the limiting case, if you put two identical products into the simulation, it will give the same total share for those two as it would if only one had been entered. It also deflates the share for any product that is similar to others in proportion to the amount of its similarity. The details are provided in the technical paper mentioned above.

The fourth model is not really a choice model, but rather a recruitment model. It estimates each respondent's likelihood of buying each concept in the simulation as though that concept had appeared in the calibrating section of the interview.

Does conjoint analysis, and ACA in particular, really work? I think so, although there's not much published evidence.

At ESOMAR in 1980, Jim McBride of Xerox and I reported a commercial study of business equipment buyers in England and Germany. Half the respondents received trade-off matrices and half computer-interactive conjoint interviews. We found that respondents' reactions to the two experiences were quite different. They were asked questions about how easy the interview was and how interesting it was. The computer interview was seen as so much easier and more interesting that there was almost no overlap in the answers. The interviews took about the same amount of time by either method, but the respondents with computer interviews tended to underestimate the time required.

Respondents in both groups were also asked their preferences among concepts that were not used in the calculation of utilities. Utilities from both groups were used to predict choices among those concepts. In both groups the prediction was significantly better than chance, but predictions were much better for the group receiving the computer interviews.

More recently, in 1986 Finkbeiner and Platz of National Analysts compared the ACA System to a "traditional" full-profile approach in a small study with six attributes. They found that the results of the two techniques were similar in terms of utilities and prediction of held-out concepts. The ACA System interview took longer, but they commented that ACA would be expected to have a timing advantage in a larger study. They also reported that ACA permitted much faster data analysis. The ACA System results were available a day or two after the last interview, while the traditional method took a week or two.

In a paper in 1982, Cattin and Wittink surveyed the commercial use of conjoint analysis at that time, and they have recently updated their survey. In a prerelease version of that paper they say,

> "We suspect that the trend is for the computer to be used more frequently for data collection. Indeed, some commercial users have predicted informally that the vast majority of projects will be conducted by computer at the end of this decade."

According to Cattin and Wittink there are several reasons: availability of the software, higher respondent interest and involvement, flexibility of custom designed interviews, ease of introducing questions to measure the consistency of the response, ability to vary the number of judgments depending on the amount of precision desired, and immediate availability of results.

Many interesting questions about conjoint analysis remain unanswered. In particular, it's not clear how best to assess validity. I'd like to see a vigorous discussion of validity at a future conference. For now, we can hope that academics and practitioners alike will be challenged to confront these questions and report their findings.

HOW TO SELL CONJOINT ANALYSIS

by
Verne B. Churchill
Market Facts, Inc.
Chicago, IL

My charge is to talk about one man's view of conjoint analysis - the climate for it and some guidelines for selling from the perspective of a research supplier. So as not to restrict this session to the ravings and prejudices of one nontechnical salesman, I prepared by talking to a number of research buyers and people from research supplier organizations about their present view of conjoint.

Despite the fact that conjoint analysis has been with us in one form or another since the early 1970's, the methodology is still not fully understood. There is a diversity of opinion about the value, shortcomings and applicability of conjoint analysis. The technique still has its enthusiasts, but it has some detractors who neither harbor uniformly positive feelings about the technique, nor understand it.

Prevailing attitudes or perceptions of conjoint suffer, at least in part, from a history of oversell. In the dim past many research buyers felt they were sold a bill of goods or were led to a set of highly unrealistic expectations for some projects involving conjoint. Conjoint was apparently presented in the early years by overzealous research salesmen as a panacea. Conjoint could not live up to those lofty expectations and, as a result, frustration resulted. A lack of confidence still persists in the minds of some buyers, much of which stems from hearsay. It's not rampant throughout the industry, but it does exist and purveyors of the service will have to contend with it from time to time. The history of our profession is filled with other techniques that came on with great fanfare, yet failed to live up to their promise.

Conjoint still is one of the most robust and valuable research products we can offer. It is a maturing, powerful tool that holds great promise for accomplishing analysis that cannot be accomplished any other way. As a profession, we've learned a great deal about what's possible through conjoint and what's not. As demonstrated at this conference, the procedures that guide our behavior are getting better and are much better than 15 years ago. Despite the developments of the past decade, which have substantially reduced the costs associated with a conjoint study, the technique remains a relatively expensive proposition. Conjoint studies are relatively visible in the eyes of corporate management, marketing management, and senior management. As Joel Huber pointed out, buyers of conjoint exhibit a special excitement and high expectations for the technique, largely because they visualize the "simulation fix" - they play games and make assumptions about changes in their brand and look at the consequences on the entire market.

The relatively high cost of conjoint in combination with this excitement and high expectations creates risks. The corporate researcher and research supplier who sell conjoint find it in their interest to proceed with great caution and special care to correctly define the objectives in achievable terms, and, in general, to do whatever is necessary to insure a satisfactory outcome.

When is conjoint appropriate? That's like asking: When do research salesmen's nostrils flare? One answer is when uncertainty exists about the optimum combination of some product's characteristics, either a new product or a redesigned product. This is the most traditional function of conjoint analysis. In the 1970's when my company introduced conjoint under the stewardship of Rich Johnson, we called it "product specification analysis." That name persisted for three or four years until the current terms became prevalent.

Another answer to the question is when it's important to determine the relative payoffs of various changes in formulation or positioning. The distinction between these two functions is that the second does not necessarily involve a search for the optimum combination of characteristics.

When else is conjoint appropriate? When the number of attributes that drive product choice in a category and, therefore, the number of possible combinations of decision options is relatively high. For instance, if there are two or three directions for a product formulation or repositioning to take, a simpler analysis, such as a concept test, may be appropriate. But if product choice is more complex and the number of options is high, then conjoint may be the only way to test the best alternative.

It helps considerably if conjoint is conducted on a product that competes in some reasonably defined category or competitive framework that includes similar brands. The power of conjoint is most obvious if the simulations can show the effects of changes on a client brand or preference shares for competitive brands.

Finally, conjoint can be regarded as an appropriate recommendation when there is reason to expect the buyer has a receptive ear. This is a pragmatic, nontechnical criterion but one that, nonetheless, should be considered. A prospective buyer who has had a previous positive experience with conjoint and has an open mind is a better candidate for a conjoint sale than a researcher or senior manager who is philosophically opposed to or distrustful of conjoint. It's a very appropriate criterion.

A similar question is: What conditions should be present to justify a recommendation for conjoint? First, the number of attributes or variables should be limited to a manageable number. Just what a manageable number is differs among practitioners. My company has a conjoint expert named Bob Ogle who is a devotee of a full-profile data collection mode. Bob Ogle would feel about Joel Huber as I feel about Don Rumsfeld or Bobby Knight. Bob Ogle's preference for the number of attributes does not exceed eight in combinations of two and four levels. There are many situations where it's necessary to go well beyond eight attributes. The industry has developed tricks, bridges and inferential techniques that theoretically allow us to deal with many more attributes. One must do that, but one must also recognize the risk of degradation of the model's predictive power when that occurs. The number of attributes is certainly an important criterion.

Second, the kinds of variables or attributes with which we deal should fit some minimum criteria. I've tried to express these with terms like describable, comprehensive, and self-evident. Describable attributes refer to the ability to communicate with respondents the product qualities or traits the researcher has in mind. Some are easy to describe and communicate, such as miles per gallon, height, width, capacity, and other physical measurements. Others become less easy or more ambiguous, such as provisions of a warranty or performance criteria for a level of gasoline. Others border on the impossible, such as freshness, appeal to children, richness, comfort, and appetite appeal. Some believe that the most difficult challenge facing the conjoint researcher is the formulation of acceptable words for this latter type of "soft" variable.

Attributes should be comprehensive, which refers to the fact that a complete trade-off should require that all qualities of a product that can be expected to drive preference should be included. They should be self-evident in meaning. The words should be obvious and as commonly understood as possible, and the attributes should be as discrete as possible. One of the most important sources of problems and error in a trade-off model is the presence of interaction. To the extent that interaction is present, the predictive power of the model is degraded, and judgments made on the basis of simulations can be potentially misleading. One way of guarding against interaction is to insure that the attributes used describe traits or qualities that differ as much as possible from one another.

Finally, consider the term "levelable," a word I coined to describe the requirement that the attribute levels selected convey the relative degrees of the presence or absence of some characteristic. For certain variables, it's a simple issue when to use inches, volume, or some other physical measure. When we're forced to use such terms, or establish a continuum ranging from "always" to "frequently" to "rarely," ambiguity is injected, because these terms mean different things to different people.

What relevance does this have to selling? The salesman has to be aware of these criteria or guidelines for attribute preparation. They form a minimum requirement for doing a coherent job of convincing the client that he should consider doing conjoint analysis with your company. The research salesman should be able to separate good and bad candidates for a conjoint study and be aware of the pitfalls.

An obvious differentiation between a good candidate for conjoint in contrast to a poor candidate is office copiers compared to a woman's fragrance. With copiers we can talk about discrete and comprehensible attributes, whereas with a fragrance the attributes involved are ambiguous and rest on the imagination of the buyer. Each Christmas we see perfume commercials that feature the Inkspots singing "I Don't Want To Set The World On Fire" with a gorgeous man emerging from a swimming pool, followed by a jet flying over the Transamerica Building. It would be a challenge to construct a set of attributes about that.

One should be wary of considering a conjoint study on products with unique qualities that depend solely on highly creative advertising, where the characteristics that drive product choice must be formed from reactions to dramatic executions or clever forms of communication. Under those circumstances the resulting utilities and simulations would lack definition.

Similarly, the products that rely largely on subjective or sensory values, rather than objective values, are less readily handled by the methodology and must be interpreted with greater care. Obviously, some qualities such as texture, interest, appetite appeal, and so on are unavoidable. When they are, use appropriate stimulus support materials, such as photographs, presentations of the real products, textures, fragrances, real colors, etc.

To whom does one sell? There is no single answer, but there are three important factors to consider: 1) previous experience with conjoint, 2) level of sophistication of research management within the client organization, and 3) job security, which recognizes that special risks are associated with conjoint studies because of their expense and greater visibility in the eyes of senior management.

The target audience for the sales presentation usually includes people beyond the research director. In a mature, sophisticated client company, senior management has likely been indoctrinated to the technique. They need not be exposed to your special magic. A company without this legacy or expertise on staff is likely to require a fair amount of stroking and cultivating.

In a number of situations, the typical sources of skepticism in the sales presentation are reversed. I've been impressed a number of times in a conjoint presentation where marketing management, the decision makers, brand management, and senior management really grasp the notion of trade-off. They immediately and enthusiastically apply it to their decision making needs. On the other hand the research director looks for flaws and challenges and becomes a barrier. This is an interesting reversal compared to most sales presentations.

What are the steps in the process of selling conjoint? First, the preparation stage requires getting your act together about the client's previous knowledge and experience, not solely with respect to research methodology, but also in terms of their willingness and ability to participate in the critical attribute preparation aspects. The early involvement of specialists is important. There are technical and philosophical disagreements among various practitioners of conjoint analysis. The research supplier company will have conjoint experts on their staff, and the client company will have experts of their own. It will greatly facilitate a successful research experience if some accord is reached early between the experts on both sides. This need not always be an adversarial relationship, and it's important that both sides feel comfortable with each other. I've been a frustrated witness to several instances where a conjoint study has proceeded to near completion, only to discover that the person on the client's side who has control over the data analysis has some substantial disagreement with the basic technique. It's in everyone's interest to get the experts together and talking as quickly as possible.

Next is the identification of attributes. If a product has had a fair amount of previous research, this will help in identifying factors that govern preference. In the absence of such a history, it will be necessary or desirable to plan a fair amount of upfront qualitative, preparatory or data reduction research to insure that the list of attributes is not only comprehensive, but utilizes proper words or terminology in the specification of levels. Like so many aspects of our craft, the likelihood of success comes down to the proper selection of words.

During this preliminary stage, one has the opportunity to test different scenarios of levels to be sure that they communicate what we want them to communicate. Finally, it is helpful to have the incorporation of the validation set.

Who should attend the sales presentation? It depends. The point has to be addressed separately for each sales situation. In general, it's desirable to include people from senior management beyond research management within the client organization. The focus of any conjoint sales presentation should be on the ultimate output - the simulation phase - the ways in which tabulated scenarios of consumer preferences can be related directly to the decision making requirements of management. Few things can stimulate an audience of marketing decision makers more than showing some illustrative or hypothetical findings of their brand fluctuating in response to different assumptions or scenarios. It's a very powerful method of selling and not just in the context of selling conjoint. Also, in the sales presentation the opportunity exists for clarifying in very specific terms the objectives of the test.

How technical should it be? As a general rule, only as much as necessary. I was in a restaurant called Ed Debevek's in Deerfield, Illinois recently. A tent card on the table showed Ed Debevek's three locations - two in Chicago and one in San Francisco. The headline read: "Ed Debevek's only minutes away, depending on where you live." How technical, then, should the presentation be? It should be very technical or not very technical, depending on who's there and who has to be convinced.

How do you explain it? First, very slowly and more importantly, through illustrations, examples and hypothetical data. In selling to people who are unsophisticated and unknowledgeable about conjoint, we follow these steps: We describe how we identify all the variables or attributes that are likely to drive choice. Then we explain the creation of levels and how they give life to the attributes. We describe the process of trade-off matrices and how through the use of an algorithm we compute the importance of each variable, without asking directly how important it is. We mention other methods of data collection, and how, after data collection, we calculate individual utilities. Then we go through simulations, first, without any changes, and then with changes in the levels of attributes.

In summary, conjoint is one of the most powerful, versatile, and strategically important research products available to serve the mission of marketing research. Sellers may encounter some lack of understanding and prejudice, but with the proper preparation, client cooperation, and sound technical support it can be sold quite readily. It's worth whatever special efforts are required, because a successfully planned and executed conjoint study brings its own special rewards.

HOW TO SELL CONJOINT ANALYSIS

by
Vincent P. Vaccarelli
Xerox Corporation
El Segundo, CA

Verne Churchill has told us about how to sell conjoint measurement. I will continue with a perspective of what happens next.

One of my advantages as an internal supplier at Xerox Corporation is that I live with my clients day in and day out. On the other hand, one of my disadvantages is that I live with my clients day in and day out. At Xerox we have four marketing research programs, including an Industry Information Library, Qualitative Research Services, Xerox Survey Center, and Concourse, which is a national panel program of 1200 members.

We have used conjoint for the past 15 years. The copier and other office product areas offer a tremendous opportunity to utilize the capabilities of conjoint measurement. We have a natural advantage. We have studied the demands for the electronic typewriter, the word processor, high volume copiers, facsimile, office information systems, and high speed data printers. We've used every technique, including fixed trade-off matrices, full-profile, and fractional factorials.

We have had up to 80 attributes, for which we did a two-stage conjoint study. The first stage had the respondents pick their attributes. The second stage sent them customized trade-off matrices.

At the peak of our experimentation, we created something called "respondent-constructed" attributes. For instance, rather than looking at speed alone for a printer, we looked at speed in conjunction with one or more applications, such as speed in conjunction with data printouts or direct mailers. We had respondents pick the importance of each element, then the system combined them. We had 185 possible attributes. Each respondent picked ten, but we could not anticipate which ten any one respondent would pick.

Now we have the Adaptive Conjoint Analysis (ACA) System from Sawtooth Software, which we have used for three studies. It is superb. It has changed the way we conduct conjoint research, as well as the way our internal clients look at conjoint measurement. We're changing the way decision makers make decisions regarding products. We're creating a discipline and a system that they feel comfortable with. When they take that conjoint interview themselves, they see conjoint in a totally new light. The machine paces them and feeds back their choices. It's a breakthrough for which I offer my compliments to Sawtooth Software.

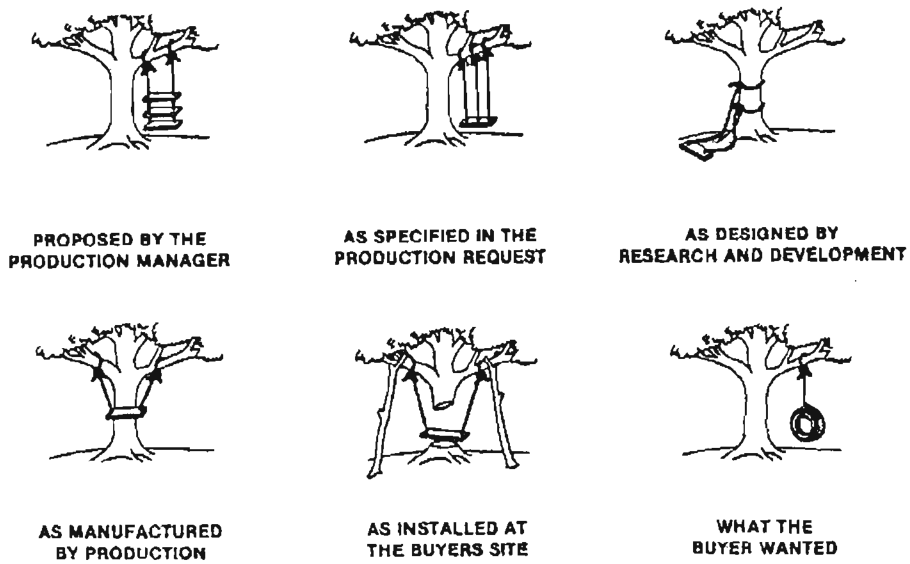
There are limitations as well as capabilities in conjoint measurement. We have identified some controlling variables to determine whether a conjoint measurement proposal and project will be successful. These are concerned with the application and its presentation to the decision maker. These are the two areas of controlling variables for guiding conjoint measurement to a successful outcome.

Recall the classic theme: Don't offer electric blankets in the summer. Similarly, conjoint measurement should not be a solution looking for a problem. It is not a new toy that we must rush out and introduce to all our information users. It should be a basic part of the research repertoire. Specifically, if you don't need it, don't use it.

How do you know when you need it and when there's a good opportunity to use it? Most important is the appropriateness of the objectives in terms of the requirements of the client. Second is communication of the value of conjoint for that objective. We use a procedure in consulting called "pacing and leading," which tunes us into the client, then we slowly lead the client to consider our recommendations. But tuning-in is crucial.

Figure 1 illustrates Xerox's major problem. We may have many features for one product. We have what some call "creeping featurism." High-tech culture offers recognition by coming up with a <u>new</u> idea or feature. There are many new features proposed daily. The key decision is how to use them to solve real customer problems.

Figure 1



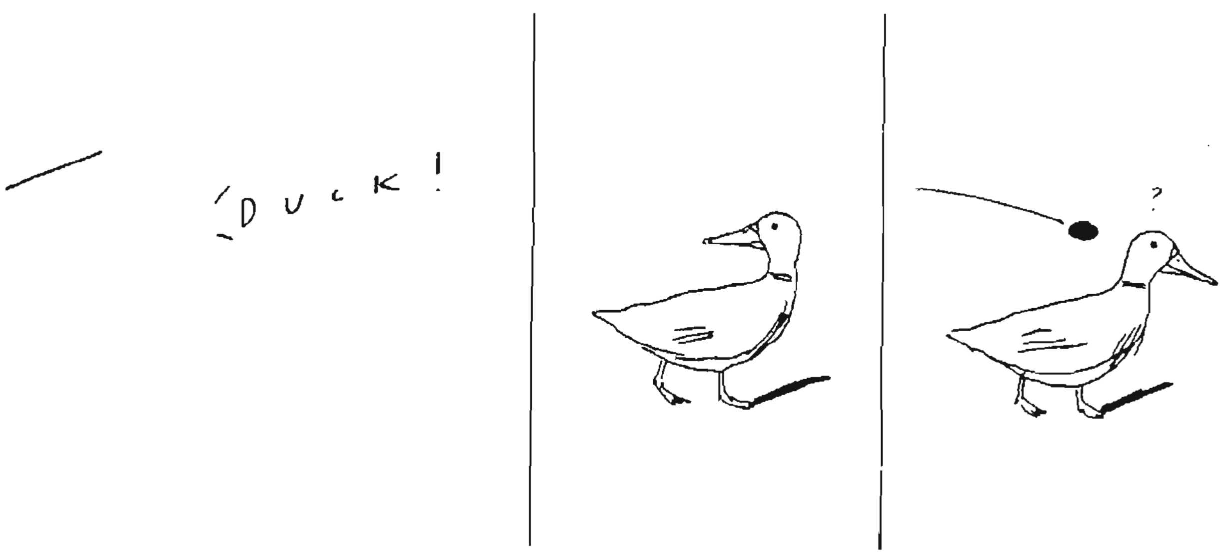
The first objective we identify for conjoint measurement is product development and customer profiles. In other words, how should we develop this copier or office-information system? Which features are most important to which customers and at what price? When we hear these questions, it's a cue for us to begin thinking conjoint measurement. At Xerox that's a very prevalent kind of question, especially in our product development groups where you can develop many products each day.

Our second conjoint objective is competitive strategy. In information system businesses, there are new products every week from many, many manufacturers. As soon as one is launched, our planners ask, "What do we do? Should we lower the price? Should we come out with another version? Should we make it more capable? What should we do now that they have just launched this product? Are we overreacting? How will this new product that's been introduced invade the preference share for the product we have out now?" The "what-if" conjoint capability is an extremely useful cue for competitive strategies.

Our third conjoint objective is imaging and positioning strategy. In reality, conjoint measurement deals with a claim that a product actually has a certain combination of attributes, which the user cannot know until he actually uses it. Therefore, we test a "claim" for a product. In that regard, we know which attributes we should get across, that is, which messages are likely to elicit which responses from the marketplace. Specifically, we identify the attributes that should be a part of the message in positioning the product. Should we position it as fast? Should we position it as economical? Should we position it as high quality?

Communication is important and certain words can miscommunicate very easily. Miscommunication can lead to some disastrous effects. With conjoint we can express the proper message with the proper words. (Figure 2)

Figure 2

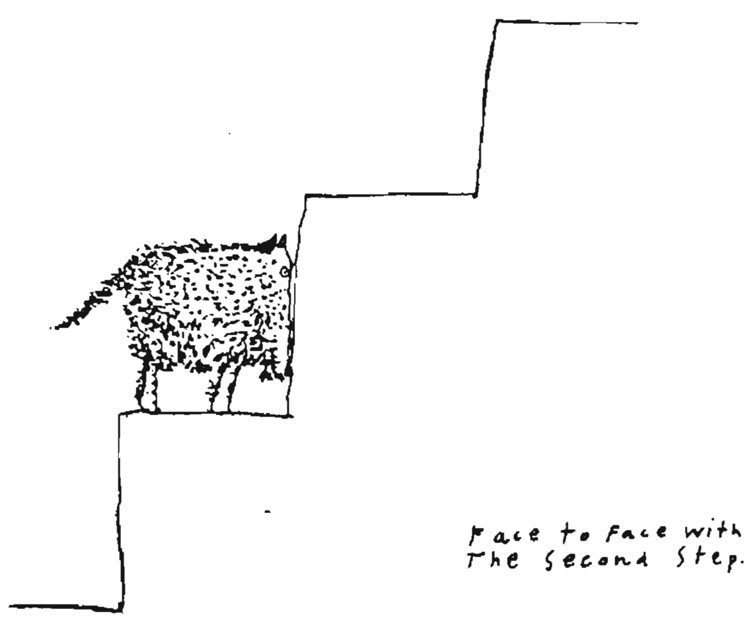


Another conjoint application area that we have just discovered involves large corporate customers who want us to build a Xerox office-information system just for them. They may have 150 users who must be happy with this system. We measure their "community utility" by having each of them take the conjoint interview that we administer at the customer's site on laptops. As a result, we can give them a very thorough proposal. It is, therefore, useful in estimation of preference among several users within one customer's site.

Also evolving as a conjoint application is decision support consultation in response to questions such as: How should we structure this decision? Very often a decision maker will come to us and say, "I just don't know how to make this decision." We first have them specify all criteria or the outcomes they might like from their decision. Then we use the outcomes as "yes/no" attributes and measure the decision-maker's utilities for each. As a result, we help them identify those decision outcomes they should trade off, and those they should not. We're using the ACA System to help structure decisions, apart from actual market research.

However, once you know when to use conjoint, you face the second step of selling it (Figure 3). We take the following steps when introducing the ACA System: First, we identify a specific client with a decision problem who may be receptive. Second, we learn the most amenable decision that might be aided by conjoint measurement. Third, we send them text and articles explaining conjoint and invite interest in a demonstration. (Not only do the text and articles introduce them to conjoint measurement, they legitimize and make clear that conjoint measurement is an established technique available for their use. Sending text and articles works very well.)

Figure 3



If they show an interest, we construct an example interview. We then do ten interviews that are relevant to their decision. We avoid talking about conjoint in the abstract. Given just these ten interviews, we conduct a private demonstration for the decision makers. However, we stress that conjoint is a support, not a replacement for executive judgment. That's extremely important. When clients see a computer making decisions that they think they are paid to make, they become nervous. You must position conjoint measurement and the ACA System as a support for the decision maker.

If they like the demonstration, we offer them a brief proposal. If the proposal is accepted, we assemble a design team of project planners who are representative of that decision maker. The design team meets two to three times a week and serves as a guiding group for the study. This team approach provides a feeling of ownership of the project and, therefore, a feeling of usefulness of the data. We give the team a prereading and demonstration of the ACA System in order to help set the objectives and outcomes. They help us develop a screener for selecting the right respondents. We help them brainstorm to rank attributes and levels in terms of their importance, insuring attributes are comprehensive, differentiating, independent, understandable, and important to the customer.

We then conduct pretests of the design during group interview sessions. After the questionnaire is administered, we ask if they think it captured their preferences (Content Validity). We then correlate their utilities to what they already have (Concurrent Validity). Then we do a concept test for Construct Validity. Given satisfactory levels of these forms of validity, we make appropriate modifications to the design and proceed to data collection.

When collecting data, we suggest continual processing and evaluation. Don't wait until you've completed 250 to 300 interviews to look at the data base. Run test simulations along the way. Merge your data as soon as possible, even daily. If there's a problem, you're likely to spot it in time.

After collecting all the data, run some test simulations PRIVATELY. Compare the results to the decision-maker's expectations and reconcile discrepancies. The decision maker may indicate that you've left out a key attribute, or double-counted one. Do this privately to allow for solutions, while preventing damage to the potential of conjoint measurement in general. Too often we over-generalize from bad experiences and under-generalize from good experiences.

Next, run the basic simulations, gather comments, and resolve issues among the design team. The design team should be exposed to the ACA System and understand how to utilize the model. Issue a report on your decisions and recommendations. Most of the audiences will not want to know just the results of the study. They prefer to know your recommendations and what data support them. The first statement in a report should read, for instance, "Increase the speed of the printer by twenty percent," followed by sample simulations. The second line might read, "Lower the price by fifteen percent," followed by supporting simulations. Recommendations followed by simulations are an effective pattern for communicating the findings.

We encourage interactive workshops. We have a number of conference rooms with televisions hooked to laptop computers. Trained facilitators bring in client teams and play "what if" scenarios. The interactive workshops best utilize the ACA System as a decision-making tool, rather than just a report generator.

In summary, the important message is to pace your client. You can neither go beyond your client, nor fall short in terms of complexity or requirements. No matter where your client is or wants to go, first work with him, then move him along. Be compatible in every possible way. That rapport not only helps with the value of a study, but also nurtures a crucial long term relationship. Conjoint should become one part of the repertoire you have available for the long term, but only when the client requires it, and with that client's understanding of how to use it.

HOW TO DESIGN A CONJOINT STUDY

by
James J. Tumbusch
Procter & Gamble
Cincinnati, OH

The market research department at Procter & Gamble has brand market research experts who advise other departments, such as product development or advertising, on market research studies. If after talking to a client they decide that a conjoint study is appropriate, they call me in as an internal consultant.

I've designed eight full conjoint studies using the ACA System in the last year, which were all successful. I had hoped to have been the first person to say, "Make sure that conjoint is appropriate," but Verne Churchill and Vince Vaccarelli have already said that. Nonetheless, it's important to ask questions to learn the objectives of the research. If conjoint is not applicable, do not recommend a conjoint study. Important questions to consider are:

- -Does the client need to evaluate a large number of concepts?

- -Is he in the dark about where to go with repositioning a product?

- -Does product development have some technical improvements that he needs to investigate to learn how consumers would react?

- -Would ordinary concept testing be appropriate?

For instance, some clients may have four, five, or six concepts they need tested. If so, conventional concept testing may be more appropriate.

Once you decide that a conjoint study is necessary, who should be involved in planning the study? You need the whole team - your product development experts, advertising brand personnel, and market research experts.

Conjoint studies, like all market research studies, need to be well designed. I was delighted with Betty Sproule's presentation about designing a perceptual mapping study. I would design a conjoint study similarly. The task requires a thorough job and the help of a market research expert. I do not consider myself a market research expert; I am a conjoint expert.

The planning process is critical and terribly difficult. We start with an initial meeting where the product development and advertising people tell us about their objectives and potential changes to the product in question. Here we learn what is on their minds and the key decision issues. We also build a preliminary attribute level list.

It is extremely useful to take the preliminary attribute level list and build an ACA System interview. For those experienced with the ACA System, this takes 20 minutes. Give the advertising, product development, and market research experts a diskette and have them take the interview; they gain a better understanding of the entire process. We generally have a preliminary meeting to develop an attribute level list, build a pilot interview, test that interview, then we meet again. The process takes at least two weeks, and it's excruciating. All this occurs before any pilot testing with consumers.

What technical issues surround building the attribute level list? First, only run a conjoint study if your real and/or hypothetical products can be decomposed into independent attributes. In the early days of conjoint, we thought only of the physical attributes of products. Today we have three distinct kinds of attributes to consider. Certainly physical attributes are still important, such as the new features of disposable diapers (a waist shield and legcuffs to prevent leakage). Another is performance benefit, such as how dry the diaper keeps the baby. Do the baby's clothes get wet overnight because of leakage from the waist? If not, that's a performance benefit. The last is psychological positioning. In the case of diapers, it's the pride that this diaper keeps the baby drier. How would you position this to the woman buying diapers? At some point you may want to involve your advertising agency in writing those positioning statements.

What attributes should be included? All actionable attributes known to consumers or feasible for change. Do not include attributes that are unlikely to change. If we were designing a study on a new version of the Ivory Soap bar, we would not include colors like yellow and blue, since Procter & Gamble will likely never change the color of Ivory Soap. Price should always be included, as should brand name. These are the two attributes that give the most technical problems in conjoint studies.

What kinds of levels should each attribute have? All levels that are appropriate to present to the consumer, including those currently available in the marketplace, plus any that technically can be changed. Levels can be discrete, such as two different scents and unscented for a hair spray product. Many attributes can be continuous, but represented by a limited number of values. Price is an obvious example of that: five dollars, ten dollars, fifteen dollars. Each attribute level must be currently available, feasible, and actionable. Do not include levels that are not actionable.

Respondents must understand each level from experience or orientation. We use solid models. For instance, we show them three different designs for a shampoo bottle. For three levels of softness for a facial tissue, we mount swatches of facial tissue on a board for them to touch. For different forms of a facial cleanser, we developed a color video to show the use of the different product forms. Just like any market research study, if the respondents do not understand what they are doing, the study will be a failure.

The levels of each attribute must be mutually exclusive for the model to work. This may not always be obvious to the users and can result in conflicting situations. For example, a facial cleanser product may be available in a three-ounce or six-ounce size. It is wrong to have a third level that includes both three and six-ounce sizes. If a respondent chooses that third level, you will not know what size he wanted.

Another problem is using words that mix degrees and levels of an attribute within levels. Terms like "removes ugly stains," "prevents toughest stains," and "scrubs away the toughest stains" confuse the issue. "Removing" stains from a fabric with a laundry product is clearly different from "preventing" stains. The terms represent different product benefits. Do not confuse the issue by having two or three levels of one, yet just one of the other. The results will be garbage. Furthermore, no interaction should occur between any two attributes. (Interaction means that the preferences for the levels of one attribute depend upon the context of the levels of another attribute.)

Another area with many problems is infeasibility. Many times you know the preliminary attribute list is loaded with infeasibilities. For example, a carbonated soft drink product may have the attributes "one calorie," "25 calories," and "100 calories." You want to know what calorie level respondents want for this soft drink. Another is "10 percent real fruit juice," "40 percent real fruit juice," and "70 percent real fruit juice." In the ACA System interview or any conjoint study, "one calorie" and "70 percent real fruit juice" are technically infeasible. Consumers will not believe such attributes. Eliminate such attributes and, when in doubt, combine them. That may be a problem in the ACA System which limits you to nine levels, but you can usually do so. An example of combining attributes involves a facial cleanser as a cream, lotion and mousse in a jar, tube, pump bottle, and aerosol can. We changed the attributes to cream in a jar, lotion in a pump bottle, and mousse in an aerosol can.

What are some non-conjoint issues? Remember that a conjoint study is just another kind of market research study. Design it fail-safe, then follow up to make sure nothing goes wrong. Design the background information ahead of time, including brand awareness, product usage, and other important aspects. Searching for structure in utilities after the fact does not work. Design a proper orientation to insure that respondents will understand what they are to do. We use story boards, color videos, and solid package mock-ups. Do some pilot interviewing with 10 to 20 respondents. Put them through the orientation, conjoint research interview, and the post-interview. This is the only way to make the study work and make appropriate changes.

As for background questions, we use the Ci2 System to ask background questions in conjunction with the ACA System. We generally ask background questions before and after the interview. The questions after the interview are questions that might bias a conjoint interview, so we ask them after the fact. The overall format for the interview is Ci2, ACA, and then Ci2.

The length of the interview rarely exceeds 20 to 30 minutes. We run six to 16 attributes with two to nine levels in a study. We have been very pleased and have had no problems with potential respondent fatigue. Occasionally a respondent cannot come to grips with a computer. Although problems can arise in any market research study, we have had no major problems in conjoint studies.

Can or should you plan the analysis in advance? Absolutely! Don't count on blindly searching for structure in the data. It does not work. Most likely you will be targeting specific consumer groups, such as, decaffeinated coffee drinkers to learn if they want a different mix of attribute levels in a coffee. Typically, we run target bases of 300 respondents, but sometimes we need as many as 450-600 to target desired subgroups. You may have to overquota target groups to obtain a representative sample.

We do central-location studies exclusively. We have not considered telephone interviewing for conjoint studies. The major reason is the need for orientation of the product features. Orienting respondents during a telephone interview is difficult. I doubt the feasibility of doing any kind of conjoint study via telephone.

DESIGNING A CONJOINT ANALYSIS STUDY

Richard D. Smallwood
Applied Decision Analysis, Inc.
Menlo Park, CA

Several discrete steps are required in the design of a conjoint analysis study:

- Identify the client's needs
- Build an analysis model
- Construct a sampling plan
- Design the interview
- Data analysis

This paper will cover each step, commenting on the major issues that arise and ways to handle them.

Client Needs

Identifying client needs is a critical step, because it defines the range of products/services that must be analyzed and determines the level of detail necessary for specifying the attributes of the product/services. Designing a study usually brings one face-to-face with an explicit trade-off between complexity and data collection costs on the one hand, and the level of detail in the product analysis on the other. To help the client make these difficult trade-offs, the market analyst must understand the client's needs for information as well as his budget limitations.

This step also requires a fair amount of discipline. The market analyst must resist the temptation to add increasing levels of detail as topical issues arise in the client organization during the study. Without this discipline, the study can become a list of the client's "hot buttons" during the study design.

Building the Analysis Model

An obvious tenet in market analysis is that one should have a strategy for using data before collecting it. Unfortunately, in many situations data are collected without any consideration of future use. This results in missing key parameters that must be estimated from other sources.

Planning a strategy for using data means that models for data analysis must be well-defined before data collection begins. In practice, the analysis model and the data collection process should be designed in parallel. The model should use data that are feasibly and credibly collected, and the data collection program should provide the precise information required by the model.

The construction of the analysis model typically involves several fundamental issues. One involves whether to describe the process by which customers decide to purchase a product vs. modeling the choice that buyers make among the available products. For example, we may want to predict how many customers will ultimately purchase cellular telephones, or we may only be interested in the choices of customers who have already decided to purchase a cellular telephone. The former situation is more difficult and requires a special set of conjoint trade-off questions different from those used to distinguish among products.

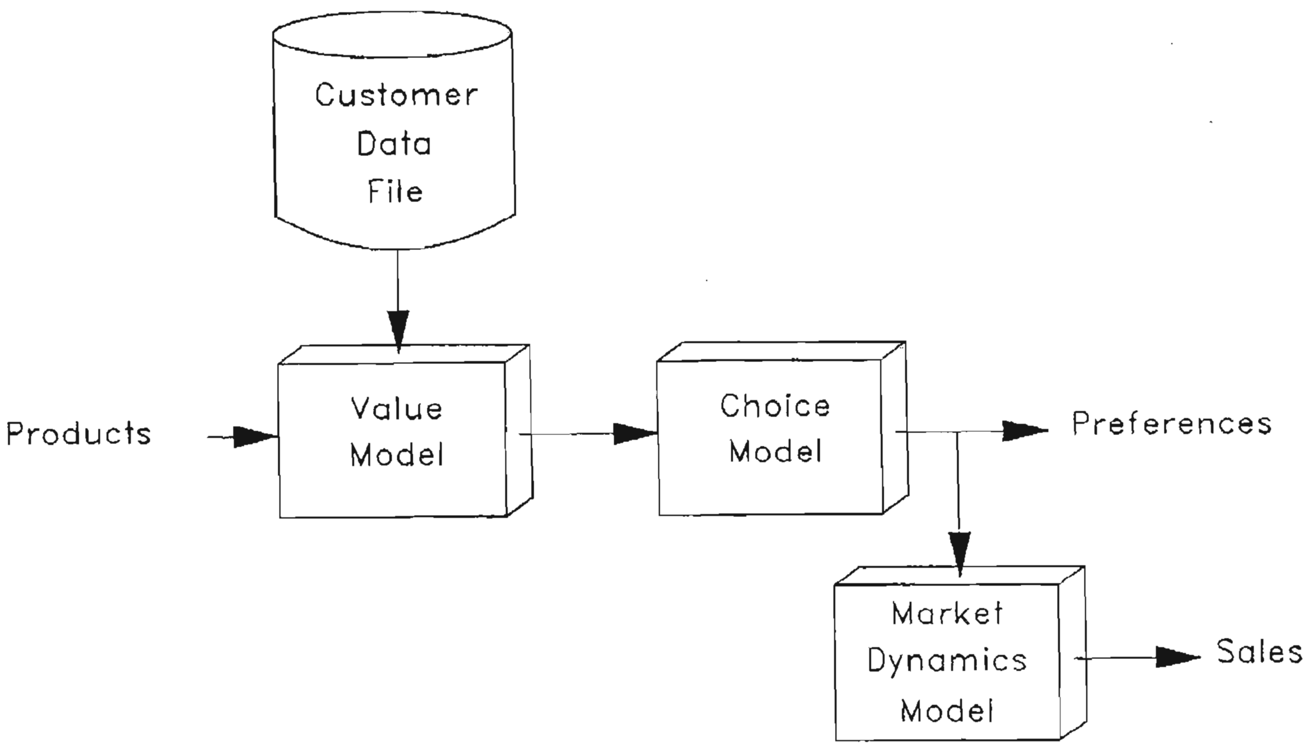
A second modeling issue concerns the representation of the customer choice process itself. Do customers typically make a two-stage choice in which they first choose among various subsets of products, then select a product from the first-stage subset? Or do they choose from the complete range of available products? The characterization of customer choices at each stage of a hierarchical choice model can have a major impact on the conjoint questions asked during the interview.

A third area of concern is product similarity. This is the so-called "red bus/blue bus" problem in which it may be necessary to account for product similarity to avoid similar products garnering more than their appropriate share within the market model. There are several mechanisms, ranging from very simple to numerically complex, for compensating for this problem.

A final concern for the designer of market models concerns the distinction between customer preferences and product sales as illustrated in Figure 1. If the market model is only designed to estimate customer preferences for products, then the modeling task is relatively easy. If, on the other hand, it is important to forecast actual sales of products over time, then a market dynamics model must be included to characterize how customer preferences are converted into the acquisition of products through various marketing actions. The market dynamics model typically includes considerations such as product awareness, reluctance to change, sales force coverage and effectiveness, and differential effects of distribution channels. Having a market dynamics model in the market analysis usually requires additional trade-off questions in the study.

Figure 1

MARKET MODEL



Constructing the Sampling Plan

Construction of the sampling plan requires some a priori description of market segmentation and the level of detail required in the information obtained from the study. Fine-grained segmentation of the market will require a careful design to insure that the proper number of samples is obtained from each market segment. This may require some stratification of samples.

There are four key issues to consider in the design of the sampling plan:

-The cost of estimation errors arising from insufficient numbers of samples

-The cost of samples

-Heterogeneity in the market

-The precision of the measurements.

A relatively simple theory determines the optimal number of samples in each market segment for a priori estimates for each of the factors. Unfortunately, most clients are unwilling to specify the first and third items on the list, so rules-of-thumb abound in practice. Although 50 samples per market segment is commonly used as an heuristic criterion for sample size, there is evidence that a smaller number of samples is sufficient in many cases.

The ultimate heuristic for determining sample size is to divide the fieldwork budget by the cost per sample. I wish I could say that this never occurs, but it often does.

Designing the Interview

In designing the interview, first determine the format. For conjoint analysis the format is determined by the medium and the location. The two most common media are personal computers and paper-and-pencil. As for location, the interview possibilities are a central facility, the respondent's office, a focus group, telephone-mail-telephone, or a mailer. The computer is usually used at a central facility or in the respondent's office, whereas paper-and-pencil is used at the last three locations. Choosing among these formats involves a trade-off between the flexibility and adaptability of computer-based interviews vs. the low cost of paper-and-pencil interviews.

The next step is to define a set of attributes and corresponding attribute levels. There are two criteria for including an attribute in a conjoint study:

-Importance to customers

-Product differentiation

If the attribute is unimportant or if all products have the same level of the attribute, it should not be included. Attributes should not overlap; they should not measure two aspects of the same part of the product (e.g., durability and reliability).

The attributes should also be easy to communicate. This can create a problem if the client is concerned about a technical attribute that is unfamiliar to most respondents. In this case, some intermediary model must be used to translate technical parameters into attributes that are familiar to respondents.

Finally, the number of attributes should be kept as small as possible, allowing respondents to trade off as large a fraction of the total attributes as possible.

Another important issue in the development of attributes is discrete versus continuous levels. Discrete attributes (e.g., brand) are those with a limited number of levels, whereas continuous attributes (e.g., price) may have an infinite number of levels. In practice, continuous attributes are sometimes converted to discrete attributes by limiting them to a small number of levels.

In many cases, a set of discrete levels for a continuous attribute is not sufficient. A good example of this is price. If respondents are to trade off attributes of an automobile against price, a reasonable range of prices might be $6,000 - $15,000. Yet trading off the value of a sunroof involves price differentials of just a few hundred dollars. The only alternative to using price as a continuous attribute is to have a large number of discrete levels. Thus, it is usually desirable to treat price as a continuous attribute. This creates special requirements for the design of the trade-off questions, as well as special demands for the analysis of the data.

Designing the Trade-Off Questions

In general, it is important to keep the number of attributes presented in each trade-off question as small as possible. As more attributes are added, respondents tend to focus on only two or three. This creates confusion, inconsistencies, and poor estimates for the utilities. On the other hand, if there are dependencies between attributes, then all interdependent attributes must be shown together in a trade-off question, with the interdependencies incorporated into the calculation of the utilities. Dependencies between attributes may not be apparent until after the data are collected. This "chicken and egg" problem is generally solved with a liberal dose of intuitive judgment; subjective assessments are used to decide what dependencies to assume in the design of the trade-off questions.

When price is traded off against another attribute, problems can arise if the price differential is not great enough. Respondents will select the alternative with the higher cost, which can cause higher utilities for higher prices. Explaining such results to a client presents a problem. One solution is to use the respondent's answers to earlier price trade-offs to adjust the price differentials shown in succeeding questions. Thus, price differentials can be made to correspond to the respondent's willingness to pay.

An important feature of computer-based conjoint measurement is the paring of the attribute list. There are several ways to do this. One is to construct a bare bones version of the product by setting each attribute at its least preferred level. Respondents are then asked which attribute they would improve first. Successive iterations of this question provide a rank ordering of attribute preferences that is better than the standard regret questions. Another technique for rank ordering attributes is to trade each attribute off against price, then use the preliminary estimates of attribute utilities to identify their relative importance.

It is difficult to decide how many trade-off questions to ask each respondent. In general, the number of questions should be large enough to give good estimates of the attribute utilities. However, the number of trade-offs should not be so large that the length of the interview becomes burdensome and fatiguing to the respondent. We generally fix the number of trade-offs per attribute at five times the number of free parameters.

If it is necessary to compensate for product similarity, a special set of trade-off questions must be asked to provide estimates of the parameters associated with the product similarity part of the choice model. In general, compensating for product similarity implies that some of the trade-off questions must involve three or more alternatives. Research is currently underway to determine how to design these three-choice trade-off questions and estimate the parameters of the product similarity model.

Data Processing

Several issues emerge in processing conjoint data. First, the choice between treating a continuous attribute as truly continuous or quantizing it into discrete levels will affect both the number of trade-off questions and the data processing strategy. If an attribute such as price is assumed to be continuous, then one or more functional forms for the utility must be postulated and tested. Specific forms used for the utility of price include linear, quadratic, and logarithmic. In most cases, a linear utility function for price is sufficient.

A second issue concerns the ordering of attribute utilities when a natural ordering for the levels of the attribute occurs. For example, if the levels of the reliability attribute are "high," "medium," and "low," then the utility for high reliability should be greater than that for medium reliability, which in turn should be greater than that for low reliability. However, if the attribute is not very important to the respondent, then the utilities may not fall into their natural ordering. It is often desirable to impose constraints on the utilities so that they reflect the natural ordering. Often this has no major effect on the fit of the utilities to the data; it may even improve the fit.

A final issue is the aggregation of individual utilities. Simple arithmetic averages are not appropriate for aggregating utilities because respondents with - "infinity" for their utility can have a disastrous effect on the average. Preference preserving aggregation schemes for individual utilities produce reasonable results. Aggregate utilities have a major disadvantage; they obscure individual correlations between preferences for attributes and can produce totally erroneous results. For this reason, we prefer to deal with individual utilities and to aggregate at the end of the product preference calculations.

Summary and Conclusions

The design and construction of a conjoint analysis study involve the careful consideration of many judgments, trade-offs, and common sense. No matter how sophisticated the tools for presenting the interview to the respondent, the design of the interview requires a special set of skills and extensive experience with previous studies.

It is essential that the client's needs, the modeling logic, and the implied data requirements be identified early to guide the planning, design, and testing of the conjoint study. The issues and techniques discussed in this paper provide some guidelines.

CONJOINT EXPERIMENTS: ANALYSIS AND INTERPRETATION OF RESULTS

by
Marshall G. Greenberg
National Analysts
Division of Booz-Allen & Hamilton, Inc.
Philadelphia, PA

My presentation will focus on the role of conjoint experiments in the research process - how the results are analyzed, interpreted, and used. In particular, I want to address a number of issues surrounding alternative approaches to the analysis of conjoint data and the use of the resulting utilities in modeling choice behavior in the marketplace.

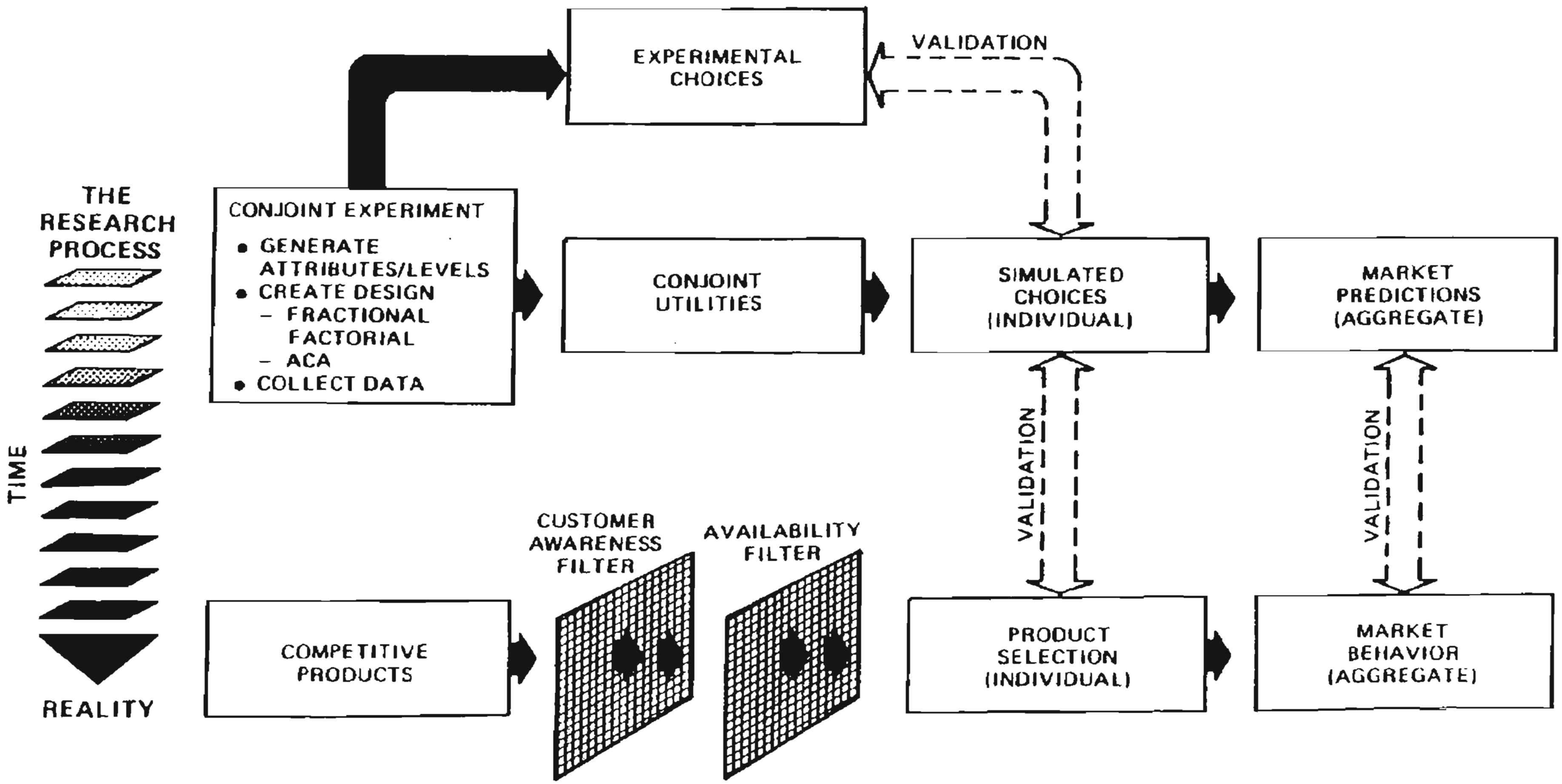
First, it is important to recognize what we hope to learn about the real world when we conduct a conjoint experiment or any experiment. Figure 1 represents a conceptualization of the research and modeling process incorporated in the typical conjoint experiment with utilities feeding into a market simulator.

As the central line of Figure 1 indicates, the research process begins with a conjoint experiment. We generate attributes and levels of attributes, create a design (either fractional factorial or an ACA System-type questionnaire), then collect data. We often supplement the conjoint task with a choice task among some individual product descriptions that can be used later for validation of the modeling process.

Generating the conjoint utilities constitutes the next phase. We take the results from our experiment, use them to generate conjoint utilities, and feed those utilities into a conjoint simulator, which attempts to simulate individual choices as we would expect them to occur in the marketplace.

The final phase is a prediction of aggregate market data, in which we attempt to estimate market share or volume under a variety of alternative market scenarios.

FIGURE 1
CONJOINT SIMULATIONS AS A MARKET MODELING PROCESS



Ultimately, the success of our predictions is tied to our ability to relate the results of the conjoint experiment to what is occurring in the real world. The bottom level of Figure 1 outlines the critical elements in the real-world marketplace. We have a set of competitive products in the marketplace, usually at a time later than the experiment that we're conducting. For any given individual purchaser, those products must pass through a couple of filters.

The first is a customer awareness filter that acknowledges that many customers are unaware of all product choices available to them at any given time. The second is an availability filter, which reflects that even among products a customer is aware of, not all products may be available. For example, several years ago Coor's beer had a high level of awareness nationwide, but it could not be readily purchased in the East.

After passing through those filters, an individual is confronted with a set of choices from among those products he is aware of which are available to him. The individual product selection process is translated into market behavior by aggregating the decisions made by all individuals constituting the relevant market.

The vertical time line on the left side of Figure 1 puts the research process on the top and reality on the bottom. The length of time between the two can be highly variable depending on the product category under investigation. This time frame has major implications for the predictive validity of the research. Research on a new product with a five-year development lead time, for example, involves making market predictions well into the future. Obviously, there are considerable risks in doing that.

To complete the conceptualization, validation is often, but not always, part of the research process. Validation can take place in many ways. Some have been discussed by other speakers in relation to the validity of perceptual mapping and conjoint analysis.

The dotted validation arrows in Figure 1 represent some possible check points for validation. For example, one can validate the aggregate market predictions coming from the conjoint experiment against actual aggregate market behavior. That would probably be the strongest form of validation, since we are most interested in predicting market behavior. However, it's a rather stringent criterion given all the variables not incorporated in the conjoint experiment that can impact market behavior.

An alternative might be to validate simulated individual choices against individual product selections. In other words, a satisfactory validation might compare what each individual does in the marketplace with what the conjoint simulator says he ought to do. If the model predicts that an individual ought to buy Brand "X" from among the available choices, then that individual ought to be buying Brand "X." One measure of validation is whether he does. Still another measure of validation involves an aggregate comparison of simulated choices against a hold-out sample of experimental choices.

The framework outlined in Figure 1 will be used to focus on a set of issues surrounding the typical uses of conjoint experiments in marketing research.

In conducting a conjoint experiment, the process of data analysis and market modeling usually involves several discrete phases:

-Generation and Interpretation of Utilities

-Choice Simulations

-Market Forecasts

-Validation

Generation and Interpretation of Utilities

Numerous methodological issues must be addressed in generating and interpreting the utilities that emerge from a conjoint experiment.

1. Type of Model Main effects models are most commonly used in commercial applications. Interaction effects models, sometimes employed by academicians, are rarely used in commercial applications for reasons mentioned by previous speakers. Additive models are very commonly used, while other models, such as multiplicative models, are rarely used. One needs to make a conscious decision about the type of model used to fit the data, because applications do arise in which the blind use of an additive main effects model will almost certainly fail to fit the data well.

2. Statistical Estimation Procedures Metric regression is the dominant procedure used for fitting respondent ratings. Based on a survey of users of conjoint analysis, Cattin and Wittink report that more than 50 percent use metric regression, with the remainder spread across a considerable

number of alternatives. A very simple technique equivalent to metric regression employs an orthogonal design and uses the mean rating score on each attribute level as the respondent's utility for that level. MONANOVA or monotone regression is probably the second most frequently used model.

3. Data Smoothing Procedures Suppose that an individual's utilities for various price levels do not form a monotonically decreasing function; that is, as price increases, the individual's utility increases rather than decreases. What do you do? Are you willing to accept a market simulator that says, "I can increase my share or my volume in this particular category by increasing my price?" Under some conditions this is acceptable, but for the most part, it would be totally implausible.

Under such circumstances you must decide whether you are dealing with respondent error or carelessness, or an error in the model. If it's respondent error, procedures exist that can be used to smooth the data by insisting that no matter what that respondent tells you, you will not assign him a higher utility for a higher price. You can assume and create appropriate monotonicity on selected attributes to solve that problem.

4. Data "Purification" Procedures What do you do with respondents who do not give consistent data? One solution is to terminate them by retaining for analysis only those respondents who provide data that can be fit by your model with a high R squared. There is some justification for doing that. If the model fits a respondent badly, it might be inappropriate to retain him in a database used for making market forecasts. Of course, remember to keep track of the number of respondents eliminated in order to have some sense of the portion of the population to which your final predictions apply.

5. Scaling/Rescaling of Utilities Scaling and rescaling of utilities can be done in various ways which impact the choice simulations. The ACA System uses zero to 100 as the scale. National Analysts typically uses the metric of the original response scale that rated the conjoint profiles, usually a 1-8 or 1-10 scale. Still others center the scale at zero and employ both positive and negative utilities.

6. Methods of Interpolation Piecewise linear estimation is commonly used for estimating the utilities of levels of continuous variables (such as price) in between the levels built into the conjoint design. Quadratic functions can also be fit, but are rarely used.

7. Attribute Importances Anyone who has used conjoint analysis is familiar with this issue. To calculate attribute importances in the usual way (i.e., based on the range of utilities from the highest to the lowest level of each attribute) it is important to do the calculations at the individual level, then average across individuals. Do not calculate average importances on aggregate data. Aggregate data will mask the importance of attributes on which people fail to agree about the rank ordering of attribute levels.

Choice Simulations

Having developed a set of conjoint utilities, the next step is to simulate the choices people would be expected to make among a set of hypothetical products, some of which are different from those described in the conjoint task. You want to create a set of predicted choices, generally (but not necessarily) at the individual level. A number of choice simulators exist for doing this. Discrete choice models are typically employed. There are four major models which I will describe in order of their historical introduction.

1. Maximum Utility Models The first is the maximum utility (MU) model, which says that each individual chooses the "product" for which he has the highest utility. "Product" refers not to a real product, but to a simulated product, i.e., an attribute profile that describes a product. The advantages of the MU model are: (1) simplicity in both concept and execution; (2) invariant results under linear transformations of the utilities (if you perform a linear transformation, the distribution of simulated choices from an MU choice simulator would be unaffected); and (3) no assumptions about the interdependencies or similarities among the choice alternatives in the simulated marketplace.

The main disadvantage stems from the MU model being deterministic, rather than probabilistic at the individual level. The model predicts that an individual will or will not choose a product, rather than assigning the individual some probability of choosing the product. To illustrate why this is a problem, consider a market in which two products are identical, except for a one-cent price difference. Using an MU model, everyone will be predicted to choose the lower-priced product. There is no allowance for the possibility that choices might split 60%-40% or 55%-45%. If one product dominates another on a single attribute (given identical configurations on all other attributes), all the choices will accrue to the dominant product no matter how small the difference. The model, therefore, tends to generate predictions of choices with exaggerated percentages at the extremes. At the low end will be products predicted with near zero shares, and at the high end products with much larger predicted shares than normally occur in the marketplace.

2. Bradley-Terry-Luce Models Historically, the second type of model employed to circumvent the deterministic nature of the MU Models was the Bradley-Terry-Luce (BTL) models. In these models the probability of an individual's choosing a particular product equals its utility as a proportion of the sum of the utilities of all the products in the simulated marketplace.

The advantages of the BTL model are: (1) mathematical simplicity and appeal; (2) probabilistic nature; and (3) a tendency to level extreme share predictions.

The primary disadvantage is that the results are not invariant under a linear transformation of the utilities. Share predictions will change if the utilities are rescaled by adding a constant to each. Nothing is inherent in those utilities or the scale from which they were derived that anchors them in reality. Generally, if you were to add one point to each of the utilities, you would want the predictions of the effects on the marketplace to remain changed. However, with the BTL model, those predictions would change.

The model also assumes "independence from irrelevant alternatives" (IIA), which refers to the similarity among choice alternatives. Basically, the IIA assumption states that the ratio of the choice probabilities for any two products is independent of the number of products or types of other products introduced into the marketplace. This assumption can lead to a kind of reductio ad absurdum. The classic example is the "red bus/blue bus" dilemma. Suppose a community has three modes of transportation: automobile, subway, and bus, and the buses are all red. A new product is introduced - blue buses that differ only in color from the red buses. The BTL model will predict that the blue buses will draw market shares not only from the red buses, but also from subways and automobiles. Yet an effective model would predict that all the blue-bus ridership should be derived from the red-bus ridership and none from other modes. Models such as BTL are not compatible with the commonly held belief among marketers that new products tend to attract customers from existing similar products.

3. <u>Multinomial Logit Models</u> The next model historically is multinomial logit (MNL), which is similar in concept to the BTL model. MNL is the result obtained when a BTL model is applied to an exponential transformation of the utilities. Its advantages are the same as the BTL model, except that it has a more formal basis in statistics, making it more mathematically appealing.

However, while MNL yields choice predictions that are invariant under an additive transformation of the utilities, multiplying all the utilities by a constant changes the predictions. Simply changing the unit of measurement on the utilities alters the predicted choice probabilities, an unsatisfactory result.

The MNL model also makes the IIA assumption, a distinct disadvantage.

4. <u>Multinomial Probit Models</u> The fourth more recently developed model is called multinomial probit (MNP). With this class of models, a product's predicted share is the proportion of a multinomial distribution where that product's utility exceeds that of any other product.

The advantages are: (1) its probabilistic nature; (2) invariant results under any linear transformation of the utilities; and (3) no assumption of IIA. The MNP model explicitly incorporates the correlations among products into its predictions, with any new alternative predicted to draw share from existing alternatives as a function of its similarities to each. A product being introduced into a niche in the marketplace will derive most of its predicted share from other products in that niche, rather than from the market leader as the BTL and MNL models would predict. In the red bus/blue bus situation, for example, the blue buses would be predicted to take all their market share from red buses, rather than from subways or automobiles.

The major disadvantage of MNP had been that in its classic form it was nonusable because one needed to estimate the correlations among alternative products, which cannot be done at the individual level. That problem can be solved, however, by dealing with subgroups or segments of the data base. Another problem was that the MNP model was computationally intractable until recently.

Carl Finkbeiner, my colleague at National Analysts, has developed and programmed estimation procedures for the MNP model along with software to incorporate the estimations into a market simulator.

Market Forecasts

Generating market forecasts is another critical step in the modeling process. Choice simulations can be converted into market forecasts by aggregating across individuals and incorporating proper adjustments for market factors. Never treat the output from the conjoint simulator as a forecast without at least considering the need to make adjustments for awareness and availability (see Figure 1). The data collection and modeling process assumes that everyone in the relevant market is aware of all products and has them available for purchase. To the extent that these conditions do not apply in the marketplace (and they rarely do), unadjusted forecasts will not accurately predict market behavior.

Validation

Both internal and external validation need to be considered. A number of criteria can be employed in assessing the validity of the conjoint analysis and modeling process.

One form of internal validation might compare the predictions derived from one set of data with another set obtained in the same experiment. For example, choices in the holdout sample, collected in the same space and time and from the same respondents, could be compared with predicted choices from the simulator. Lack of validity can derive from respondent error or from model misspecification. Either or both can occur, leading to poor correlations between simulated and actual choices.

External validations against the marketplace also are subject to both respondent errors and model misspecification. Furthermore, because time, space, and participants differ between the experiment and the marketplace, other sources of error may impact on validation measures. The actual product may be perceived differently than the concept tested in the experiment. The environment may change with time. If you're predicting three to four years into the future, for instance, the price of gasoline can change dramatically, affecting peoples' utilities for fuel economy in automobiles. New attributes or levels can be introduced into the marketplace.

In conclusion, I've conceptualized the relationships among the conjoint experiment, methods of data analysis, choice simulation models, and what happens in the marketplace itself. Awareness of these relationships, as summarized in Figure 1, can help to make the proper decisions during the research process by focusing on the realities of the marketplace and the behavior we ultimately are trying to predict and understand.

ANALYSIS AND INTERPRETATION OF CONJOINT RESULTS

by
Donald Marshall
Smith Kline & French Laboratories
Philadelphia, PA

The company I work for, Smith Kline & French Laboratories, is a manufacturer of prescription pharmaceutical products that we sell to physicians. We use conjoint analysis regularly to forecast the market potential for new products - both ours and our competitors'. Conjoint analysis works quite well for forecasting in the pharmaceutical market, since physicians have all been extensively trained to make rational decisions in selecting a drug for the treatment of a particular illness.

In addition to the obvious reasons for forecasting new products, we also use conjoint to provide an indication of which features will be most critical to the product's success. We forecast the sales potential of competitive new products, not just to estimate their potential, but also to see from which existing products they draw, their competitive advantages, and where they might be vulnerable. Because of the nature of the pharmaceutical market with its seven to ten year research and development lead time and rigid regulatory environment, we cannot take full advantage of the flexibility of conjoint analysis to help design or position products.

At Smith Kline we have been using conjoint analysis in one form or another for 10 years. Initially we used the full profile technique. Respondents were given a number of cards, and each card described a hypothetical product profile. The respondents sorted these cards to rank order them in terms of their overall desirability. The data were analyzed by the MONANOVA program to calculate utility scores for each feature.

While this technique worked satisfactorily, we experienced some difficulties. These problems were primarily associated with the need to keep the number of features and levels low enough so that the physicians did not have an unmanageable number of profiles to rank order.

About seven years ago we changed our approach when we switched to a computerized trade-off system that was a previous incarnation of the ACA System. This computerized approach gave us the flexibility to include a large number of features and levels without losing respondent interest.

Additionally, this new approach used least squares regression to calculate the attribute utilities instead of MONANOVA. Theoretically, this is appealing, since MONANOVA is not guaranteed to yield the global optimum, and it has been reported to converge on local optima in some studies (Wittink and Cattin, Journal of Marketing Research, Feb. 1981). From a practical point of view, this made little difference, since the stress scores we had achieved with MONANOVA were so small, that even if we were at a local optimum, little room for improvement would have occurred by moving to the global optimum.

After several years of conducting computerized trade-off studies, we have found that a sample or cell size of 50 usable respondents is sufficient to provide good, reliable results. Increasing the cell size beyond 50 is unlikely to have any meaningful effect on the model results.

We almost always analyze the data at the individual respondent level, rather than analyzing data that have been aggregated across respondents or into groups of respondents. Analyzing aggregate data only allows you to determine average responses without capturing information relevant to individual variation or idiosyncrasies. The use of individual data permits those variations to be captured and modeled. Yet we occasionally must aggregate data, particularly when using a variation of conjoint that is designed specifically for pricing analyses. The type of questioning involved in these studies suffers from a high level of respondent fatigue, and is best analyzed by randomizing the question sequence and aggregating data for the analysis.

Our first step in the analysis of the data is to determine which data to include in the analysis and which to exclude as being of little or no practical value. Largely this involves determining which respondents we want to exclude because their data were too inconsistent to be of value. While we have used several different measures of consistency in the past, we have recently used the correlation between the predicted and actual answers to the calibration concept questions that are asked at the end of the ACA System interviews.

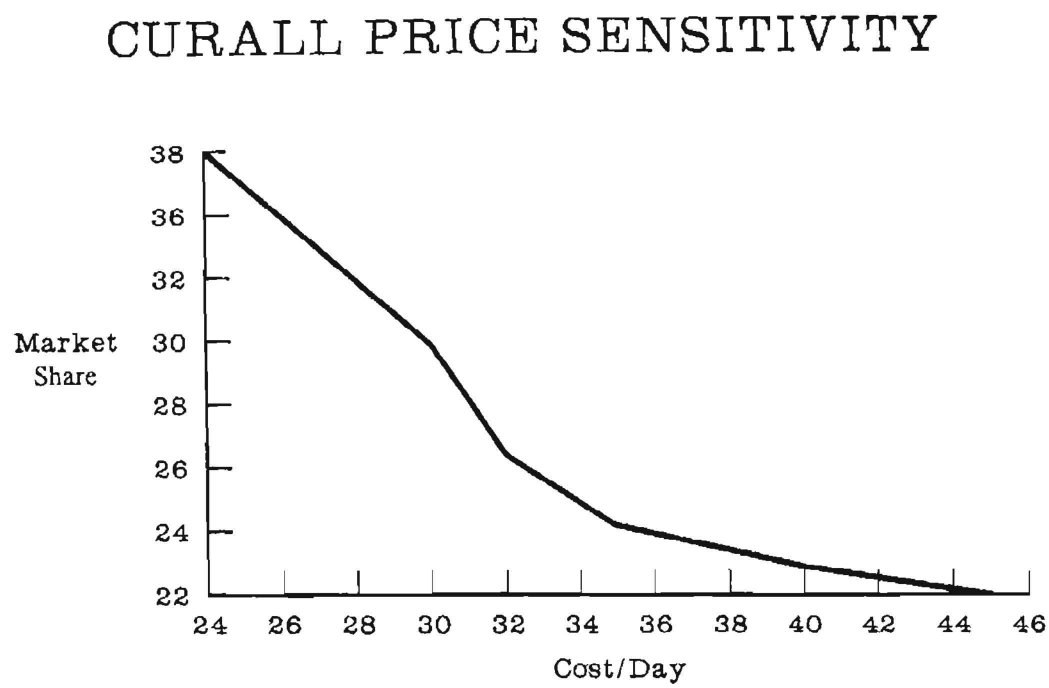
No hard-and-fast rules exist for deciding which correlations identify "consistent" data. We generally use a correlation of 0.85 as the minimum acceptable correlation for data gathered amid the noise and confusion of a medical convention, which is where most of our data are collected. For data collected in a location without the distractions of a convention, we set a minimum correlation of 0.9 to 0.95. We discard 15% to 20% of the data collected in a convention setting. Data collected in a marketing research or other similar facility results in a discard rate of about 5%, even with a tighter consistency requirement.

Once utilities have been calculated, we look to see what they tell us. We ask: "Do the utility scores make sense?" The first step involves checking the utility scores for an attribute to insure that the obvious preferences have been maintained in the utility scores. In the pharmaceutical market, this means checking that a greater degree of efficacy has a higher utility or is preferred to a lower degree of efficacy, or that few side effects are preferred to more side effects. Most of the features that we deal with have an obvious rank order that can be assigned to their various levels, although some features such as manufacturer have no obvious rank ordering. In this case, we review the utilities to see if they are "reasonable."

Next, we look at the utilities for each attribute to see which is most important to this market. Importance is measured by the ability of one attribute to contribute to the product-use decision. The attribute that contributes most to swinging this decision is the most important attribute. The ability of one attribute to contribute to the decision is indicated by taking the difference between the utilities of the best and worst levels of that attribute. Obviously, the attribute with the greatest range in utilities can contribute more utility points than the other attributes and is, therefore, the most important overall.

Figure 1 shows the range of utilities for eight attributes of a hypothetical antibiotic market. The two attributes with the greatest range in utilities are nephrotoxicity (kidney failure) with a range of 136 and ototoxicity (hearing loss) with a range of 123. Doctors, of course, have a strong preference for an antibiotic that causes neither of these side effects.

Figure 1



The next most important feature in this market is efficacy against a type of bacteria known as Pseudomonas. Continuing this sort of analysis through the five remaining attributes will yield a rank order of the relative importance of each attribute. This importance ranking is only relevant to this particular market and to these eight attributes. If any of these attributes were measured at a level that does not currently exist, we no longer are measuring what is important to the current market, but rather what will be important in the future.

There is a school of thought that takes this analysis one step farther and calculates a measure called determinance or importance, which measures what percent of the product use decision is accounted for by each attribute. This is done by summing the utility ranges for all the attributes to obtain a market range in utilities. Dividing the range in utilities for each attribute by the market range in utilities yields the percent of the decision controlled by that attribute.

For instance, in Figure 1 adding 100 + 71 + 62 and so-on yields a market utility range of 657. Dividing the anti-Pseudomonas efficacy utility range of 100 by the market utility range of 657 indicates that anti-Pseudomonal activity accounts for 15.2% of the decision to use a product in this market.

While we develop these determinance or importance measures, we don't spend a great deal of time on them, since they indicate what is important to the market as a whole, rather than what is important to individual products in that market. The critical features of one product are often of minor importance to the market. In this example (Figure 1), dosing is the least important attribute in the market. Despite this, a significant dosing advantage could provide enough additional utility to give an otherwise average product a significant competitive advantage. For this reason, we prefer to evaluate what is important to a particular product.

If the utility scores appear reasonable, we proceed to running market simulations. First, we simulate the current market. After constructing a profile of all of the products in the market on each of the features studied, we run the simulation. Whenever possible, these profiles are based on physicians' perceptions or images, rather than actual profiles as supported by clinical data. Since there are frequently large differences between the actual clinical profiles of a product and its perceived profile, this can make a substantial difference in the model predictions.

When running simulations, we always weight the respondents based on their use of products in the market of interest to us. This results in doctors who are heavy users of products receiving greater emphasis or importance than doctors who are light users. Additionally, we try to weight the respondents so that the various specialties are proportionately represented.

Once the weightings and profiles are satisfactory, we run the simulation and compare the market shares predicted by the model to the actual shares reported in the audits. If the predicted shares are reasonably close, we proceed with our analysis. If the predicted shares are not reasonably close, we review the model and data to determine the reason for the discrepancy.

The first step is to review the attributes to insure that no relevant attributes have been omitted. If the attributes appear correct, we next consider the various marketing variables that could affect the market share. This includes such variables as the amount of promotional effort supporting the different products; the quality and believability of that promotion; and the length of time that each product has been on the market. In the pharmaceutical market the order of entry is a significant factor in a product's performance. Generally, a second product that is very similar to the first will achieve about half the sales of the first product. For discrepancies that seem reasonable in light of these marketing variables, we adjust the model predictions with external factors, which range from 0.5 to 2.0.

After the model simulates the current market, we simulate the potential for the new product of interest. We start by profiling the new product on each of the attributes included in the study. Since new products have usually been in development for 7 to 10 years, we generally have a good idea of the "most likely" product profile.

Then we determine an appropriate external factor for the new product, which involves finding an analogous product and using its external factor as a starting point. This external factor is adjusted appropriately based on the amount of promotion it's expected to receive, its quality and believability, and a penalty for being late on the market.

The simulation is run, and the resulting market share predictions indicate the "most likely" share for the new product when it reaches maturity. If the available data leave ambiguity about the product's profile, the simulation is rerun after changing the new product's profile as often as necessary to cover these ambiguities.

The simulation results indicate what the new product's "most likely" market share should be at maturity, if the product is perceived as it was profiled, and if no other changes occur in the market. By comparing these share predictions with the current shares, you can determine how vulnerable each existing product is to the new product, and how much of the new product's growth comes from each existing product, assuming there is no market growth. This information helps in developing our competitive strategy, since it helps target our promotion against the most vulnerable products. On the other hand, if the new product will be competitive with one of our products, this shows our vulnerability and weakest points, enabling us to develop a defensive strategy.

The results may need to be interpreted carefully, depending on the model used. The share of preference model with the adjustment for product similarities tends to overestimate the new product's impact on very similar products, and underestimate the new product's impact on less similar products. In other words, the adjustment indicates that too much of the new product's growth will come from very similar products.

Not using the adjustment, however, will underestimate the new product growth, estimated to come from very similar products. Thus, we use both models to develop an estimate of the upper and lower limits of potential losses to the new product.

Given the nature of the FDA regulations governing the pharmaceutical industry, we have very little control over most of the attributes that profile a product. We generally include 9 to 12 attributes in a study, but price is the only attribute that we can directly control. Therefore, we take a very close look at the new product's price sensitivity. The estimation of a product's price sensitivity is done by running the simulation several different times, with the only difference between the runs being the price at which the new product is profiled. The most likely price should be near the middle of the price range studied. The model results tell us how changing the new product's price is expected to change its market share at maturity.

From these data we construct a price sensitivity curve such as the one shown in Figure 2 for the hypothetical antibiotic called Curall. This graph clearly shows that increasing Curall's price should result in a dramatic loss of market share. This knowledge of price sensitivity then helps to develop a pricing strategy that supports the corporation's market objective for Curall, whether that objective is to maximize share, revenue, or something else. If the objective is to maximize revenue, these share predictions must be converted into revenues. In the pharmaceutical market this is a simple calculation that considers such things as the discontinuance rate, non-compliance rate, and average length of therapy, in addition to the number of patients in the category.

While the calculations are relatively simple, some of the numbers needed to perform these calculations can be difficult to develop. This analysis can be taken one step farther by estimating profitability at the different prices.

Figure 2

CURALL MARKET
Relative Feature Importance

Feature	Range in Utilities
Pseudomonas Efficacy	100
Enterobacter Efficacy	71
Klebsiella Efficacy	62
Ototoxicity	123
Nephrotoxicity	136
Pain	54
Dosing	38
Cost	73

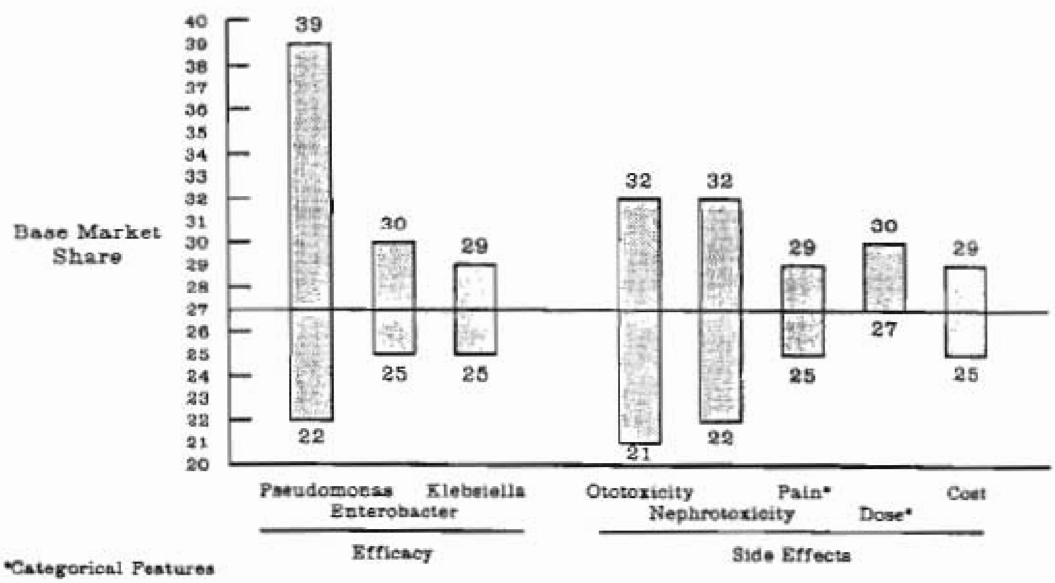
While price is usually the only variable that we can control, it is useful to evaluate the new product's sensitivity to changes in each of the other attributes included in the study. This sensitivity analysis is performed by altering the new product's profile on one feature at a time, running the simulation for each change, and recording the change in the predicted share for the product. In order to fairly evaluate which attribute the product is most sensitive to, the attributes should be changed by a constant amount for each run. We use +10% for our sensitivity runs, although this is purely arbitrary. Some of the features in the study will be categorical in nature, like dose, and a 10% increment or decrement make no sense. Instead, change these attributes to the next appropriate level. For example, a twice-a-day dose could be improved to a once-a-day, or lowered to four-times-a-day.

The results of this sensitivity analysis can then be presented in a graph such as Figure 3, which clearly shows how much could be gained or lost from a constant change in each feature. The horizontal line in this graph represents the share that was predicted for Curall with its most likely profile. The bars above that line indicate the potential share point gains from a 10% improvement in the product profile, while the bars below that line indicate the potential share point losses if the profile is made worse.

Figure 3

CURALL SENSITIVITY ANALYSIS

Share Pt. Gain (Loss) from 10% Increase (Decrease) in Each Feature



This type of sensitivity analysis provides much more useful information than the overall attribute importance measures described earlier, since this analysis indicates what is important to the product of interest. It is not uncommon to find that the most important feature to the market is not the most important feature to a particular product. Sensitivity analysis indicates what is important to a product based on the utility scores and the product profiles, whereas the market importance analysis only considers the utility scores. Thus, the sensitivity analysis makes allowances for the competitive environment, but the importance analysis does not.

Additionally, if the sensitivity analysis is done early enough in the product's development, it can provide useful input to the product development process by indicating the critical attributes of the product. Developmental work can focus on those critical areas to insure these benefits are maximized.

In closing, I would like to emphasize that the model predicts the new products market share at maturity, and also assumes that the product's image is consistent with the profile in the model. It is frequently necessary to extend this analysis to determine the product's share during the time needed to reach maturity, about five to seven years in our markets. Additionally, once the product is introduced, its developing image should be closely monitored to see if any discrepancies are developing between this image and the profile used in the model.

PRESENTING RESULTS

by
Peter B. Bogda
M/A/R/C Inc.
Dallas, TX

While some people enjoy thumbing through the latest edition of the American Statistician, most people who settle in for a conjoint presentation have two concerns:

-Will I be able to use these results?

-Will I still be awake when this thing is over?

Getting down to business, not statistics, should be the theme of a conjoint presentation. Direct your content to the lowest level of expertise in the room. Of course, you may have to adjust that by "importance weights" of the individuals there. It's all right to talk over the head of a junior analyst, but you better not talk over the head of the brand manager or senior research personnel. The goal of your presentation is to promote understanding, get down to business, and elicit action. You need to know who your audience is, and their experience, interest, and familiarity with conjoint. Then gear the technical aspects to that level.

Being a generalist and a businessman, I always have my marketing science people present to answer technical questions. The success and understanding of a conjoint presentation is based on an acceptance of the method. A client or member of the audience needs to understand the basic components and results.

Part of gaining acceptance is presenting reasons for choosing conjoint. What advantages does it have over other methods? How does it fit with the objectives of this study? Present the objectives in conjoint terms. For instance, use a trade-off analogy, such as "it's not the best pizza in town, but they deliver." Does delivery compensate for less than an ideal product? Or "it's a great car, but it's not worth the money." Are people willing to pay for superior quality?

Provide a brief example of the conjoint task and compute a simple form of utility from the example, so that people who have never been exposed to conjoint will gain a basic understanding. Figure 1 shows an example of a simple trade-off rank order and individual brand preference. This individual prefers your brand and is not price sensitive. In the left-hand column is your brand, and in the right-hand column your competitor. The prices are $1.00 and $2.00. The choices are your brand at $1.00, your brand at $2.00, your competitor at $1.00, and your competitor at $2.00.

Figure 1

UNDERSTANDING:

Compute a set of simple utilities

Converting rank order into "magnitude" (i.e., lowest rank = highest value) and averaging for the level (i.e., marginal values) is the simplest form of utility.

	Your Brand	Competitor	
$1.00	Rank 1 = 4 points	Rank 3 = 2 points	Average = 3.0 pts.
$2.00	Rank 2 = 3 points	Rank 4 = 1 point	Average = 2.0 pts.
	Avg. = 3.5 pts.	Avg. = 1.5 pts.	

Next, compute a simple set of utilities. Convert rank order - 1,2,3,4 - to some order of magnitude. Explain to the audience what the computer does. You can simplify this by saying that your brand has an average of 3.5 points. The price of $1.00 has an average of 3.0, and $2.00 an average of 2.0 (Figure 2). Remembering that conjoint is an additive model, illustrate what the results do when you add them. Your brand at $1.00 had a utility of 6.5, but at $2.00 it was 5.5, etc. The modeled rank was 1, 2, 3, 4, and in this example the true rank also was 1, 2, 3, 4. But with real data, the true rank may differ.

Figure 2

UNDERSTANDING:

ILLUSTRATE THE RESULTS OF THE ADDITIVE MODEL:

	Modeled Rank	"True" Rank
Your brand at $1.00 = 3.5 + 3.0 = 6.5	1	1
Your brand at $2.00 = 3.5 + 2.0 = 5.5	2	2
Competitor at $1.00 = 1.5 + 3.0 = 4.5	3	3
Competitor at $2.00 = 1.5 + 2.0 = 3.5	4	4

Explain who was included in the sample. Was it children, a quota sample, or a random sample? Present the task; show the attributes and levels used. Report the technique - was it full-profile, pair-wise, or self-explicated?

We've been doing conjoint studies for about two years. At M/A/R/C we have a product called CASEMAP, which stands for "conjunctive approach to self-explicated multiplicative additive preference." CASEMAP was developed by Seenu Srinivasan of Stanford and Gordon Wyner of M/A/R/C, and has goals similar to those of the ACA System.

In a representative study, the sample was 400 female heads of household. The attributes and their levels were: Brand A, B, and C; a glass or plastic package; and price of $0.50 or $1.50. The technique was full-profile.

Marsh Greenberg explained that utility values can be converted to a familiar scale. Ever since Bo Derek, everyone knows the meaning of "10," and it might be a good scale to use. Price trade-offs can be illustrated by calibrating the "gaps" between the brands or features. If you look at the utility scores for Brand B, A and C in Figure 3, Brand B has 8.0, A has 5.0, and Brand C has nothing. In Figure 4 we overlay price. One of the easiest ways of talking about trade-offs is to ask for Brand A with a utility of 5.0 what price change would need to occur to boost its utility to 8.0. The answer is to reduce your current price of $1.00 to $0.70.

Figure 3

AGGREGATE UTILITY VALUES OF BRANDS

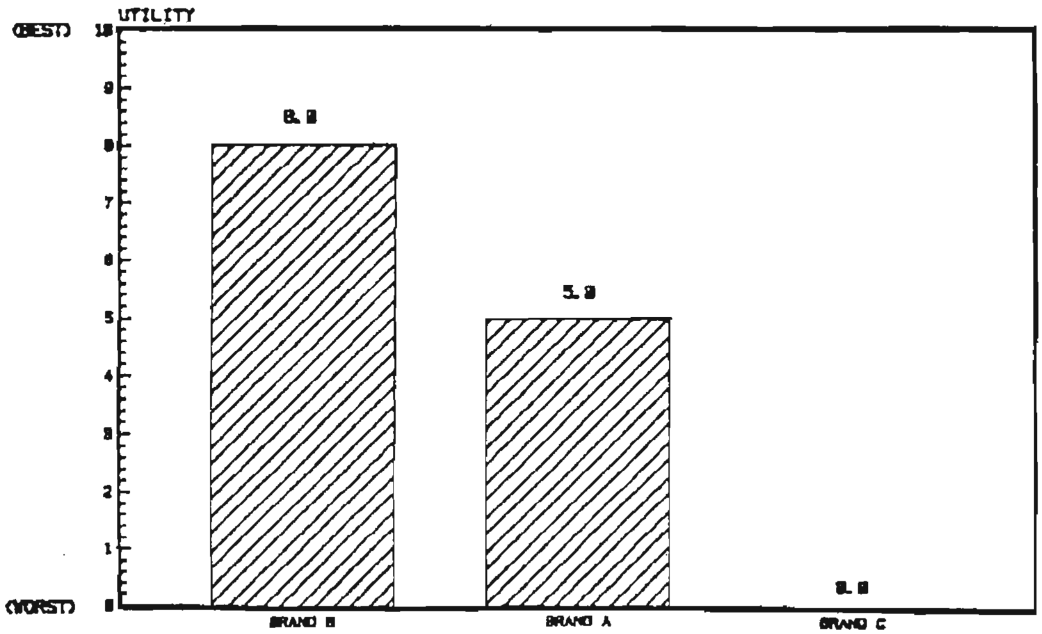
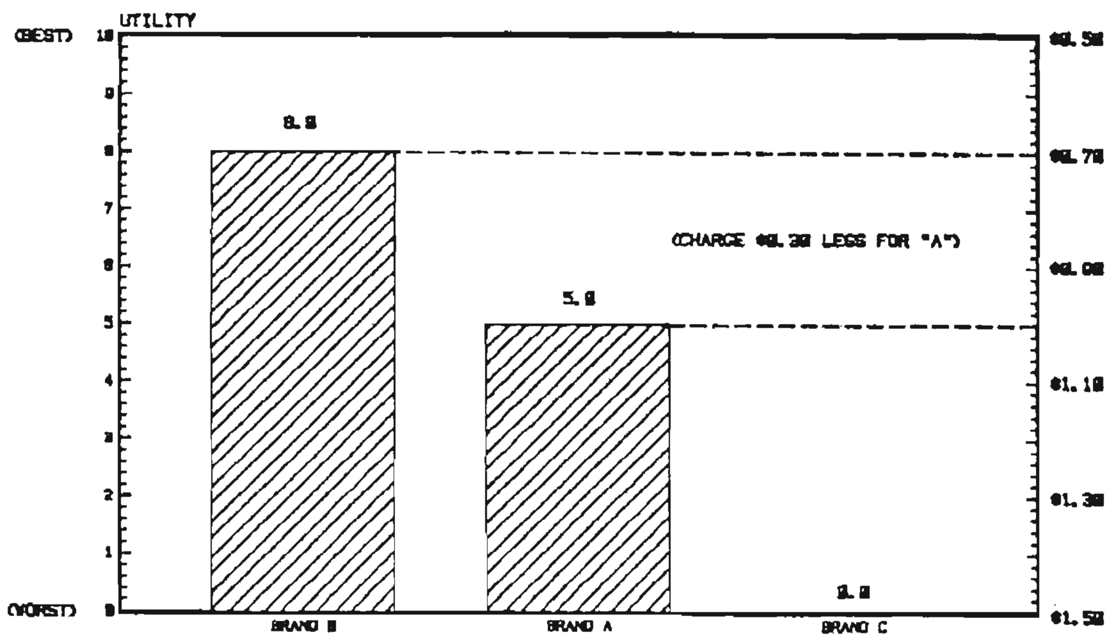


Figure 4

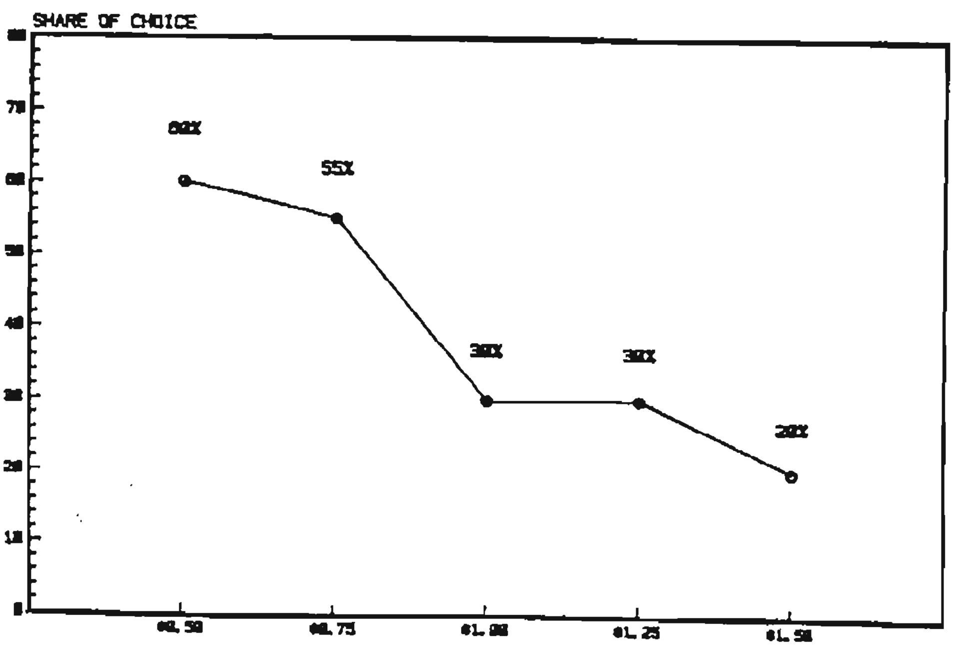
PRICE DIFFERENTIALS BETWEEN BRANDS



You also can simulate choices under various market scenarios and simulate the effect of price change on choice. Figure 5 is an example of a share-of-choice chart for presenting this data. It shows that going from $0.50 to $0.75 reduces your potential share of choice from 60 to 55, whereas going from $0.75 to $1.00 (a terribly steep slope) drops the share of choice to 30. If you go from $1.00 to $1.25 it stays the same, but it drops again out to $1.50. You can show these one at a time. There's one utility for brand, one for package, and one for price.

Figure 5

SHARE OF CHOICE AT VARIOUS PRICES



Next we calculate share simulations by individual (Figure 6). Product 1 might be Brand B at a price of a $1.50, the package is glass, and the share simulation suggests that the respondent would give 40% of his share of choice to that brand. Moving across to Brand A it's 35%, and for Brand C it's 25%.

Figure 6

INTERPRETING:

EXAMPLE OF SHARE SIMULATION:

	Product 1	Product 2	Product 3
Brand:	B	A	C
Price:	$1.50	$0.75	$0.50
Package:	Glass	Plastic	Plastic
Share of choice:	40%	35%	25%

Another way of looking at the data segments respondents into clusters based on their utility values. After segmenting, the utility values are used to predict preferences for each cluster (Figure 7). Cluster 1 prefers Brand B in a glass container and is willing to pay $1.50. That particular cluster would give 50% of their choices to that mix. This is a totally different way of looking at the data.

Figure 7

INTERPRETING:

SEGMENTATION EXAMPLE:

	Cluster 1	Cluster 2	Cluster 3
Brand preference	B	A	C
Package preference	Glass	Glass	Plastic
Price threshold	$1.50	$0.75	$0.50
% of sample	50%	20%	30%

Now I'd like to present an example. I had a client a few years ago who acquired another company that had a favorable reputation in the diet or low-calorie food business. One of my client's food technologists discovered a way to make a product that had virtually no oil. (Most the calories in this product come from the oil.) We did product testing prior to the introduction of the new company. We reduced the oil level to practically nothing - 5% or 2% - and we tested that product against the major low-calorie brands. Even at that level of oil, we showed a significant preference for this new product. If we raised the level of oil to half that in the regular low-calorie product, the product was bumping up against even a regular non-diet version. This was a wonderful product.

The firm had a name they thought they could capitalize on. We were brought in to design an approach that would help them "maximize their position." What price should they charge? What oil (calorie) level was the best? They wanted to use the new brand name of the company they just acquired. We did a full-profile study and included the major competitors in the business.

The director of marketing was at the presentation. Price was no problem. The oil content fell in a range that was profitable, and the R&D person said, "That's terrific." We did a taste claim and used the attributes "tastes as good as regular" or "tastes as good as low-calorie." There was no noise; it discriminated and provided very rich data. Then we got to the chart on brands, and their brand was practically zero. Silence.

This takes us back to the subject of any presentation: Know your business. The question haunting me was: How can I tell the CEO that he spent 50 million dollars to acquire a company, but he can't use the name? Rather than saying, "That's your problem," I recommended alternatives:

- The brand doesn't seem to fit this category, which the company had never been in. However, if you have extra marketing money and can assume the product is as good as it tested, then use marketing support dollars to promote it.

- Give it away, send out coupons, just get it in consumers hands to overcome the brand name problem.

- Advertise for another $10 million.

If none of those options is practical, I suggest adding in other data. When we did the study, we learned how many people used this product and why they were on a diet. We correlated that data, and it supported the poor fit of the brand name.

As an example, our technicians pulled together a complete array of all combinations that had been studied. We showed 50 permutations on one chart. The brand manager had to read through 30 brands out of 50 before his brand name appeared, yet it was the cheapest, lowest in calories, and had the highest taste preference.

We had a very good meeting, but the report ended up in the circular file. They used their brand name, and used the study to set a price for the product that was test marketed for a year. At the end of that year, the brand had a real share level of only 0.8%.

In conclusion, summarize your findings, talk about the benefits, and propose actionable recommendations.

PRESENTING RESULTS

by
Marc Prensky
The Boston Consulting Group
Boston, MA

I work for the Boston Consulting Group (BCG) and our business is not market research. Instead, we do "competitive business strategy formulation" at the general management level. We work for very large corporations and help them with the extremely difficult decisions they face about their investments over the medium-to-long run in a very changing and competitive environment. We help them decide how to generate a sustainable competitive advantage, i.e., a way to build a wall around their business and make money over time.

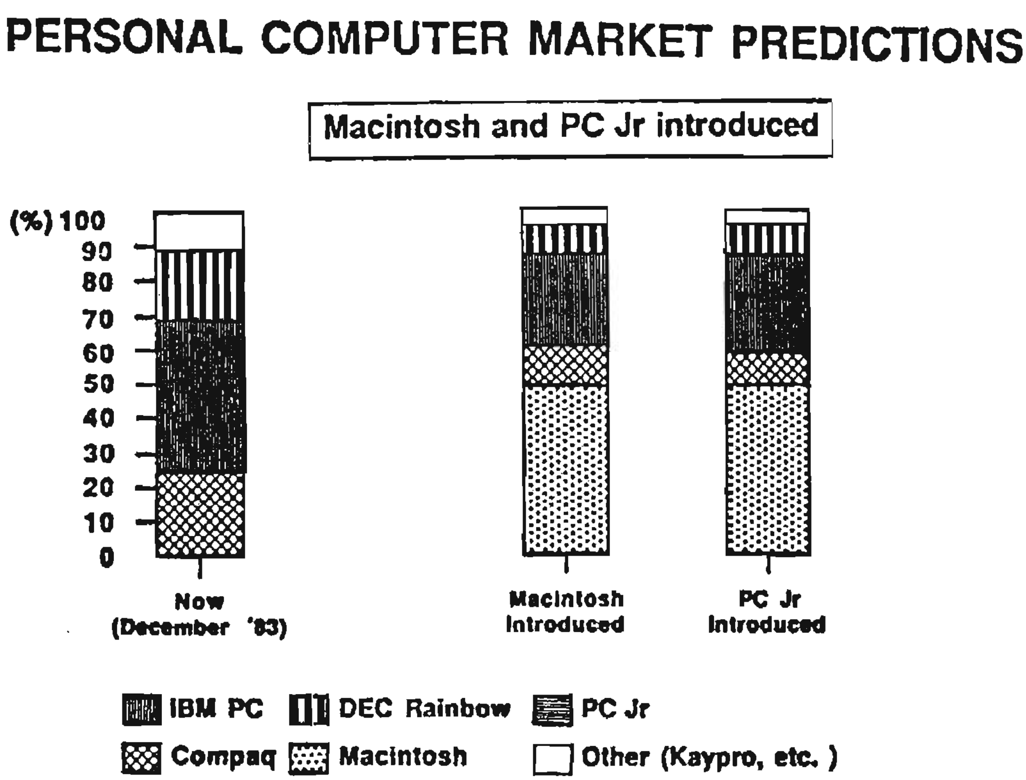
Twenty years ago, the founder of our company decided that those walls have to do with costs on the supplier's side. If you have lower costs than your competitors, you will have lower prices, and eventually get bigger market shares. The approach at that time was to become the lowest cost producer, cut prices, add capacity, and grab market share. However, it is also sometimes possible to build a competitive strategy around demand, by figuring out how to sell more products than your competitors, perhaps even at a higher price.

Conjoint analysis is one way to do demand analysis, which we at BCG find particularly useful and interesting. It's not something that we sell as a product, but rather that we use as a tool. The microcomputer has provided us with a neat way to collect and analyze data. Conjoint can be done quickly, and as you know, everyone wants his results yesterday. Once the data are collected, basic analysis takes a couple of days. Another reason we use conjoint is that the utilities plus the answers to demographic questions provide an incredibly rich database of consumer information. A clever analyst can use a program like LOTUS 1-2-3 to extract a variety of useful information from a conjoint database.

I first became interested in conjoint in 1983. We were looking at the personal computer market, which was just about to take off. The IBM PC was on the market and Apple had put out its Lisa, which later flopped. The press was talking about the PCjr from IBM and the Macintosh. Nobody had seen either of these products yet, but people were talking about them daily. The "standard wisdom" at the time was that the PCjr was going to be great, but the Macintosh would be another flop like the Lisa. BCG went to a computer show to interview some people using conjoint and (according to what we've heard today) did just about everything wrong in using the technique.

We collected data from potential computer buyers using trade-off matrices, ran it through a first choice simulator and came out with a result that simulated the current market with market shares (Figure 1, bar 1). We included Compaq, the IBM PC, the DEC Rainbow, and a number of others. IBM got half the market in the simulation, which seemed right at the time. That's an interesting start, we thought, but what happens if we introduce some other new products? We introduced the Macintosh, which we assumed was not going to do very well, but the results were totally counter intuitive. A few months later, what the model had predicted actually happened, and the technique started to look good to us, although we still thought it might be due to a fluke. (Figure 1, bar 2.)

Figure 1



We also did further competitive gaming with the model and made other predictions, which also came true. As time went on, we found that conjoint does seem to work, though I'm not always sure why. It even works despite mistakes you make in doing it, with the exception of a few key ones.

In presenting conjoint results, my goal is to keep it as nontechnical as possible. We're usually presenting to top management, who do not want to hear about a black box or jargon. Even the word "utilities" is jargon to them. We use a lot of graphics, explain weights and importance briefly, and assign anything technical (and sometimes even anything numerical) to our backup personnel.

In the explanation phase, we reinstate our objectives and explain why we are using conjoint. We usually use conjoint for very specific objectives. (For example, we may want to know the relative importance of adding one feature versus another. We may want to segment product users to predict the success of a new product, or to assess competitive reaction to an existing product.) We usually describe the sample, show how the results achieve the goals, and recommend action. This whole process is often summarized into five or ten slides in a short presentation. People in management don't have time for unnecessary details. They just want to know what's happening and be assured that it's done right. By the time you're presenting results to them, you hope that they trust you.

Our methodology can be described in three sentences. We sample the consumers by having them make pairwise trade-offs. We assign value points to each feature, and predict sample purchases. Then we extrapolate to the total population.

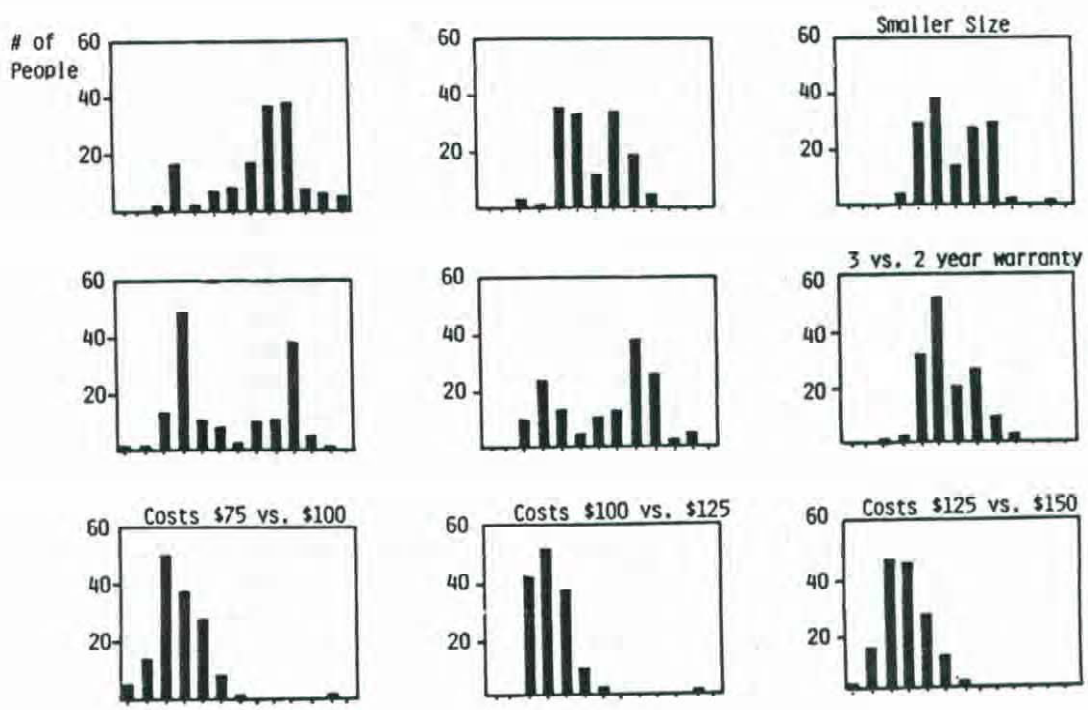
I find that graphics are almost always better than numbers when presenting findings. In fact, I sometimes try to avoid numbers entirely. Show as much as you can with simple and concise graphs. One or two graphs can show the entire results of the analysis, when directed specifically at the objectives. If possible, be a bit dramatic. For example, we had a client who held the major share of the market for a technical product. He was trying to retain his share as the market changed and competitors came in. He also was trying to segment the market by varying design features. The questions were: How do you segment? What features are important? We were able to show all the utilities scores for all the attributes (i.e., smaller size, two-year warranty, price, etc.) on a single graph (Figure 2). Most of the features have "normal" utility distributions, but some are bimodal, implying segmentation of consumer preference. Based on these results, two separate products were designed and marketed.

Figure 2

CONSUMER SEGMENTATION BY FEATURE

Value Point Distributions

(Each bar represents the number of people with a value score between 0-100, 100-200, etc.)

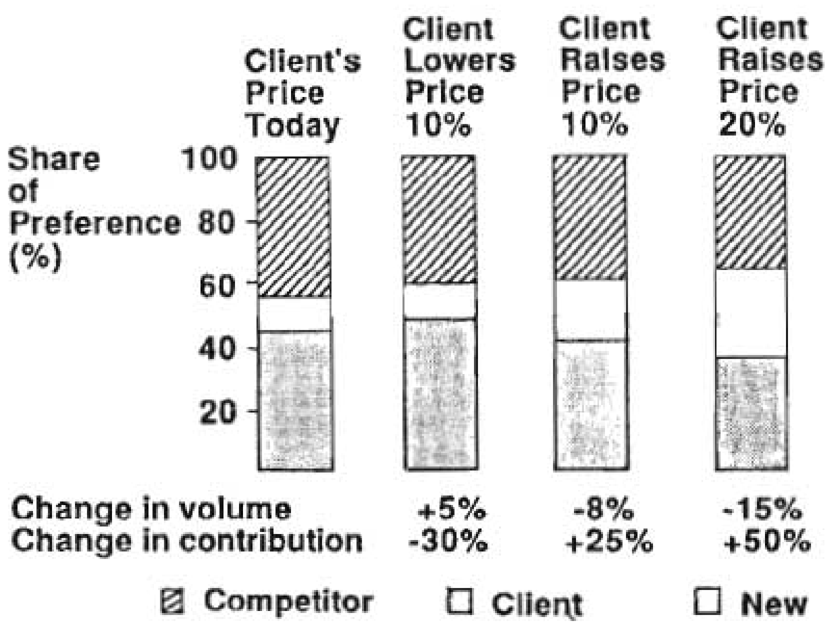


For another client, we were evaluating a new competitive entry. They asked: Is this a big deal? Should I be afraid of it or not? Conjoint cannot predict accurately what future shares will be, particularly if the product is something new. But it can predict directionally whether it will be a big or a small deal. Everything told us this was not going to be a big deal, and that eventually turned out to be true.

We also experimented with the model, and lowered the client's price, which leads to a rise in his volume (Figure 3). You also can see that as you raise prices, volume goes down. That's no surprise. Why did we do that? If you raise prices, depending on the balance between volume and costs, your total contribution may increase. Since we know a great deal about the cost and finance side, we were able to calculate the best alternative. Regardless of the entry of the new product, we recommended raising the price to maximize the contribution. This was a simple example of how we put the demand and supply side of the equation together.

Figure 3

Pricing

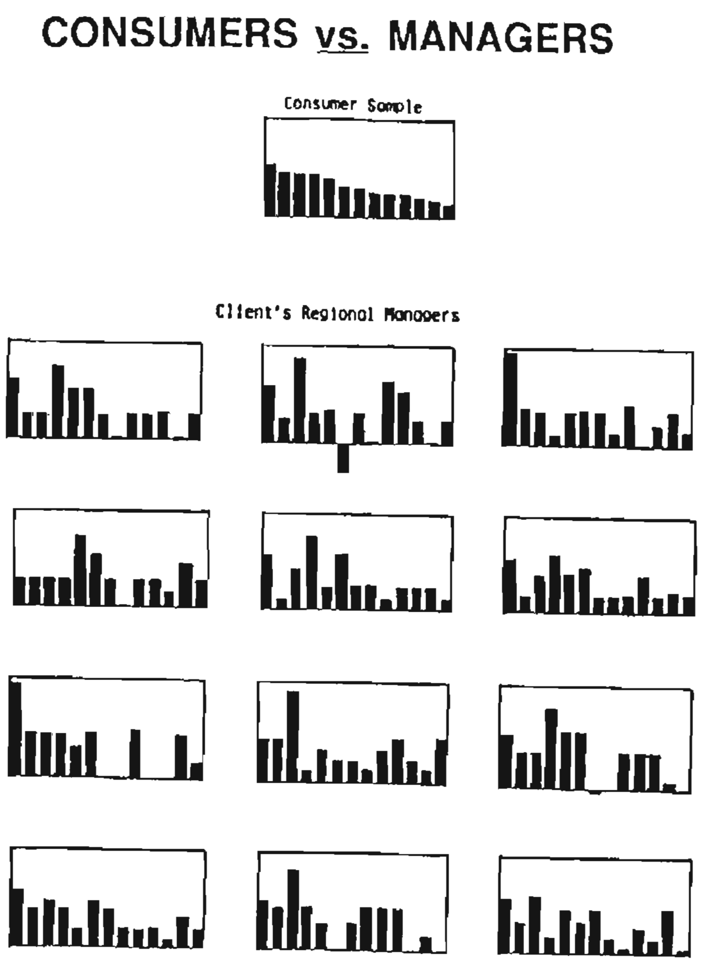


Observation

- Even with new entry, margin appears best maximized by price <u>increase</u> rather than cut

It's also interesting to involve the client management in taking the conjoint interview. Often the people who manage the product are out of sync with the marketplace. People do not easily accept being told they're out of sync during a presentation, so they need evidence. In one market study, we took the consumer utility values and averaged them for the entire marketplace. Then we gave the same questionnaire to the managers and individually plotted each of their utility weightings. The results (Figure 4) dramatically pointed out that there were a lot of differences of opinion, including differences of opinion with the marketplace.

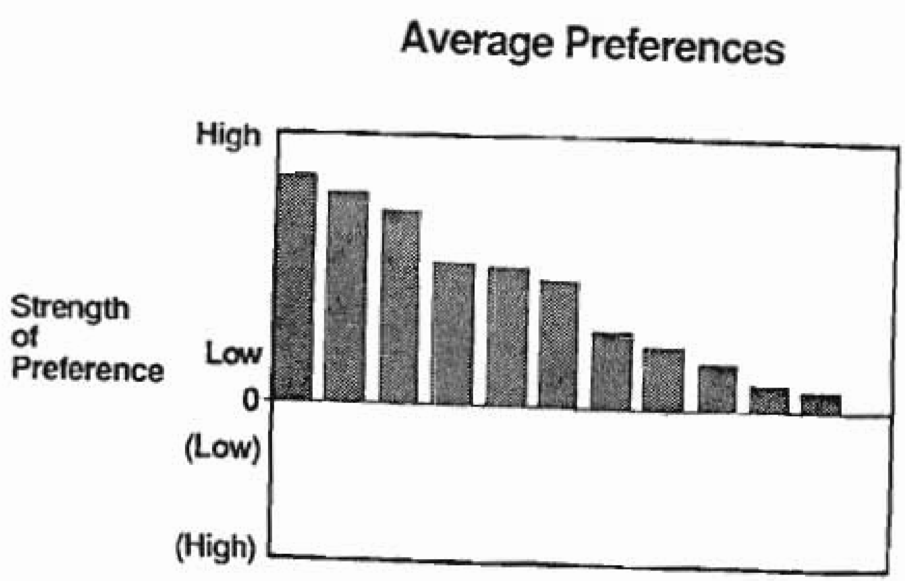
Figure 4



A few caveats for using conjoint: Many factors besides features often determine product shares and success. You can account for these by "external factor" weightings, but I do not recommend this. You've just spent time gathering people's real opinions - real data from real people telling you real things. It's directional and imperfect, but real. Whether you're using factors that worked previously or just guessing, I'd just as soon use the data as is, or guess at the whole thing.

A conjoint analysis reflects a single point in time. Things change and marketplaces can move very rapidly. Not only does the marketplace move, but people's knowledge of products moves. With new products in particular, people's attribute preferences and weightings change as they become more knowledgeable. The first time we studied a medical product, accuracy was very important. The product measured the level of something in the body, and accurate measurement was very important to anyone with this condition. But once it became generally known that all these products met a standard level of accuracy, other considerations like size or ease-of-use became more important. For certain products you have to redo conjoint regularly.

Figure 5



Conjoint is based on a number of assumptions that may not reflect the real world, yet they're as close as we can come. They may, in fact, be better for addressing a problem than the assumptions in other techniques, but they still may not be real. As an example, attributes will never be totally independent, even if they are statistically independent.

The most important caveat for conjoint or trade-off analysis is to interpret the results of conjoint directionally, rather than numerically. When somebody says, "This product gets a 14.3 share," I cringe or laugh depending on the context. You cannot do that. Conjoint does an incredibly good job of separating what's important from what's unimportant. It's very good at giving general ideas of the future success of a product. But if you treat the results as more than general, you run into problems. Figure 5 shows an actual graph with the client information removed. The scale is high/low. If you're concerned about the the last item on the right, you shouldn't be. If you're not concerned about the first item on the left, you ought to be. That type of directional message can be more powerful and usually better remembered than numbers.