Sawtooth Software

RESEARCH PAPER SERIES

History of ACA

Richard M. Johnson,
Sawtooth Software, Inc.,
2001

530 W. Fir St.
Sequim, WA 98382
(360) 681-2300
www.sawtoothsoftware.com

History of ACA

Rich Johnson, Sawtooth Software

History of ACA

Although ACA makes use of ideas that originated much earlier, the direct thread of its history began in 1969. Like much work of the development in marketing research, it began in response to a client problem that couldn't be handled with current methodology.

The Problem

In the late 60's, I was employed by Market Facts, Inc., and the client was in a durable goods business. In his company it was standard practice that whenever a new or modified product was seriously contemplated, a concept test had to be done. The client was responsible for carrying out concept tests, and he answered to a product manager who commissioned those tests. Our client's experience was like this:

> The product manager would come to him and say: "We're going to put two handles on it, it's going to produce 20 units per minute, it will weigh 30 pounds, and be green." Our client would arrange to do a test of that concept, and a few weeks later come back with the results.
>
> But before he could report them, the product manager would say: "Sorry we didn't have time to tell you about this, but instead of two handles it's going to have one and instead of 20 units per minute it will produce 22. Can you test that one in the next three weeks?" And so on.
>
> Our client found that there was never time to do the required concept tests fast enough to affect the product design cycle. So he came to us with what he considered to be an urgent problem - the need to find a way to test ***all*** future product modifications ***at once***. He wanted to able to tell the product manager, "Oh, you say it's going to have one handle, with 22 units per minute, weigh 30 pounds and be green? Well, the answer to that is 17 share points. Any other questions?"

Of course, today this is instantly recognizable as a conjoint analysis problem. But Green and Rao had not yet published their historic 1971 article, "Conjoint Measurement for Quantifying Judgmental Data" in JMR. Also, the actual problem was more difficult than indicated by the anecdote above, since the client actually had 28 product features rather than just four, with some having as many as 5 possible realizations.

Tradeoff Matrices

It seemed that one answer might lie in thinking about a product as being a collection of separate attributes, each with a specified level. This presented two immediate problems: a new ***method of questioning*** was needed to elicit information about values of attribute levels, and a new ***estimation procedure*** was needed for converting that information into "utilities."

Our solution came to be known as "Tradeoff Analysis." Although I wasn't yet aware of Luce and Tukey's work on Conjoint Measurement, that's what Tradeoff Analysis was.

To collect data, we presented respondents with a number of empty tables, each crossing the levels of two attributes, and asked respondents to rank the cells in each table in terms of their preference. We realized that not every pair of attributes could be compared, because that might lead to an enormous number of matrices to be ranked. After much consideration, we decided to pair each attribute with three others, which resulted in 42 matrices for the first study. One has to experience filling out a 5x5 tradeoff matrix before he can really understand what the respondent goes through. If the respondent must fill out 42 of them, one can only hope he remains at least partially conscious through the task.

To estimate what we now call part-worths, we came up with a non-metric regression procedure which found a set of values for each respondent which, when used to border the rows and columns of his matrices, produced element-wise sums with rank orders similar to respondents' answers.

Although we learned a lot about how to improve our technique for future applications, this first study, conducted in 1970, was a success. The client was enthusiastic about his improved ability to respond to his product manager's requests. The client company commissioned many additional tradeoff studies, and similar approaches were used in hundreds of other projects during the next several years.

In those early days there was less communication between practitioners and academics than we enjoy today. My early work at Market Facts was done almost in a vacuum, without the knowledge that a larger stream of similar development was taking place simultaneously among Paul Green and his colleagues. ACA benefited greatly from interactions with Paul in later years, and as time passed it became clear that Tradeoff Analysis was just a different variety of Conjoint Analysis. As such, it made all of the assumptions common to Conjoint Analysis, plus one more big one.

Assumptions and Difficulties

Like other conjoint methods, we assumed that the utility of a product was the sum of values attaching to its separate attribute levels. However, Tradeoff Analysis, like all more recent "partial profile" methods, further assumed that respondents' values for attribute levels did not depend on which other attributes were present in a concept description. In other words, Tradeoff Analysis required a strong "***all else equal***" assumption regarding the attributes omitted from each matrix.

This made Tradeoff Analysis uniquely vulnerable to distortion if attributes were not considered to be independent by respondents. Suppose two attributes are different in the mind of the researcher, but similar in the mind of the client, such as, say, Durability and Reliability. When trading off Durability with price, the respondent may fear he is giving up Reliability when considering a lower level of Durability. This kind of double-counting can lead to distorted measures of attribute importance.

As another example, price is often regarded as an indicator of quality. As long as partial profile concept presentations include both price and quality, one would not expect to see reversals in which higher prices are preferred to lower ones. However, if concept presentations include only price but not quality, price may be mistaken as an indicator of quality, and respondents may act as though they prefer higher prices.

Similar problems still characterize all partial profile methods today, and it remains critically important when using partial profile methods to remind respondents that the concepts compared are to be considered identical on all omitted attributes.

A second problem unique to Tradeoff Analysis was the difficulty respondents had in carrying out the ranking task. Though simple to describe, actual execution of the ranking task was beyond the capability of many respondents. We observed that many respondents simplified the task by what we called "patterned responses," which consisted of ranking the rows within the columns, or the columns within the rows, thus avoiding the more subtle within-attribute tradeoffs we were seeking. This difficulty appeared to be so severe that it motivated the next step in the evolution that resulted in ACA.

Computer-Assisted Interviewing

Researchers who began their careers in the 70's or later will never be able to appreciate the dramatic improvement of computer technology that occurred during the 50's and 60's. In the late 50's, "high speed" computers were available, but only at high cost and in a limited way. While at Procter and Gamble in the early 60's, I considered myself lucky to have access to a computer at all, but I would get one or at best two chances in a 24-hour period to submit a programming project for debugging. A single keypunch error would often render an attempt useless. It's amazing that we were able to get any work done at all under those conditions. However, in the 70's time-sharing became common, providing an enormous improvement in access to computers.

In marketing research, we depend heavily on data from survey respondents. When something is wrong in a set of results, it can often be traced to a problem at the "moment of impact," when the respondent provided the data. Originally having been trained as a psychologist, I was interested in the dynamics of what happens in interviews. When time-sharing and CRT terminals first became available, I became excited about the possibility of using them to enhance the quality of market research interviews.

I still remember an experience at Market Facts when I arranged a meeting of the company's management to demonstrate the radical idea of computer-assisted interviewing. I had borrowed the most cutting-edge CRT terminal of the time, which consisted of a tiny 3-inch screen in an enormous cabinet. I had shrouded the CRT with a cloth so I could introduce the idea of computer-assisted interviewing without distraction. The meeting went well until the unveiling, when, with a flourish, I removed the cloth to reveal the CRT. When they saw the tiny screen in the enormous cabinet, everyone in the room began to laugh. And they continued laughing until I ended the meeting. Fortunately, CRT terminals also improved rapidly, and it wasn't long before computer-assisted interviewing became entirely feasible.

Pairwise Tradeoff Analysis

Ranking cells in a matrix can be difficult for respondents, but answering simple pairwise tradeoff questions is much easier. For example, we could ask whether a respondent would prefer a $1,000 laptop weighing 7 pounds or a $2,000 laptop weighing 3 pounds.

Consider two attributes like Price and Weight, each with three levels. In a 3x3 tradeoff matrix there are 9 possible combinations of levels, or cells. We could conceivably ask as many as 36 different pairwise preference questions about those 9 cells, taken two at a time.

However, if we can assume we know the order of preference for levels ***within*** each attribute, as we probably can for price and weight, we can avoid asking many of those questions. Suppose we arrange the levels of each attribute in decreasing order of attractiveness, so that cells above and to the left should be preferred to those below or to the right.

Then we can avoid questions comparing two classes of cells. First, we can avoid questions comparing any two cells that are similar on one attribute, such as comparisons of cells in the same row or in the same column. This avoids 18 possible questions. Of the possible questions that remain, we can avoid those that compare any cell with another that is dominated on both attributes, such as below it and to its right. That eliminates another 9 questions, leaving a total of only 9 for which we cannot assume the answer.

Among those, if we are lucky (and if respondents answer without error) we may have to ask only two questions to infer the answers of the remaining seven. For example, in the 3x3 matrix with lettered cells,

a	**b**	**c**
d	**e**	**f**
g	**h**	**i**

if we were to learn that **c** is preferred to **d** and **f** is preferred to **g**, then we could infer that rows dominate columns in importance, and we could infer the rank order of all 9 cells. Likewise, learning that **g** is preferred to **b** and **h** is preferred to **c** would permit inference that column differences are all more important than any row differences, and we could also infer the entire rank order.

By the mid-70's, computer technology had advanced sufficiently that it became feasible to do computer-assisted Tradeoff Analysis using pairwise questioning. A large project was undertaken for a U.S. military service branch to study various recruiting incentives. The respondents were young males who had graduated from high school but not college. A large number of possible incentives were to be studied, and we were concerned that the required number of tradeoff matrices would strain the capabilities of our respondents.

My associate at Market Facts, Frank Goode, studied strategies for asking pairwise questions that would be maximally informative, and wrote a question-selecting program that could be used to administer a pairwise tradeoff interview. We purchased what was then described as a "minicomputer," which meant that it filled only a small room rather than a large one. Respondents sat at CRT terminals at interviewing sites around the U.S., connected to a central computer by phone lines. Each respondent was first asked for within-attribute preferences, permitting all attributes subsequently to be regarded as ordered, and then he was asked a series of intelligently chosen pairwise tradeoff questions.

We found that questioning format to be dramatically easier for respondents than filling out tradeoff matrices. The data turned out to be of high quality and the study was judged a complete success. That study marked the beginning of the end for the tradeoff matrix.

Microcomputer Interviewing

In the late 70's, Curt Jones and I founded the John Morton Company, a partnership with the goal of applying emerging analytic techniques in a strategic marketing consulting practice. We were still utterly dependent on the quality of the data provided by respondent interviews, and that led to many problems; but I remained convinced that computer-assisted interviewing held at least part of the answer.

By that time the first microcomputers were becoming available, and it seemed that computer-assisted interviewing might finally become cost-effective. We purchased an Apple II and I began trying to produce software for a practical and effective computer-assisted tradeoff interview. Use of microcomputers meant not having to be connected by phone lines and not having to wait for one's turn in time sharing, and also provided powerful computational resources. My initial approach differed from the previous one in several ways:

> First, it made more sense to choose questions that would reduce uncertainty in the part-worths being estimated, rather than choosing questions to predict how respondents might fill out tradeoff matrices. This was a truly liberating realization, which greatly simplified the whole approach.
>
> Second, it made sense to update the estimates of part-worths after each answer. Each update took a second or two, but respondents appeared to appreciate the way the computer homed in on their values. One respondent memorably likened the interview to a chess game where he made a move, the computer made a move, etc.

> Third, a “front-end” section was added to the interview, during which respondents chose subsets of attributes that were most salient to them personally, as well as indicating the relative importance of each attribute. The questioning sequence borrowed some ideas from “Simalto,” a technique developed by John Greene at Rank Xerox. We used this information to reduce the number of attribute levels to be taken into the paired-comparison section of the interview, as well as to generate an initial set of self-explicated part-worths which could be used to start the paired-comparison section of the interview.
>
> Finally, those paired-comparison questions were asked using a graded scale, from “strongly prefer left” to “strongly prefer right.” Initially we had used only binary answers, but found additional information could be captured by the intensity scale.

Small computers were still rare, so the experience of being interviewed had considerable entertainment value. We found that an effective way to sell research projects was to pre-program a conjoint interview for a prospective client's product category and take an Apple with us on the sales call. Once a marketing executive had taken the interview and had seen his own part-worths as revealed by the computer, he often couldn't wait to use the same technology in a project.

We purchased several dozen Apple computers, and began a fascinating adventure of using them all over the world, in many languages and in product categories of almost every description. Those early Apples were much less reliable than current-day computers. I could talk for hours about difficulties we encountered, but the Apples worked well enough to provide a substantial advance in the quality of data we collected.

ACA

In 1982, I retired as a marketing research practitioner, moved to Sun Valley, Idaho, and soon started Sawtooth Software, Inc. I had been fascinated by the application of small computers in the collection and analysis of marketing research data, and was now able to concentrate on that activity

IBM had introduced their first PC in the early '80s, and it seemed clear that the “IBM-compatible” standard would become dominant, so we moved from the Apple II platform to the IBM DOS operating system. With that move we achieved 80 characters per line rather than 40, color rather than monochrome, and a large improvement in hardware reliability.

ACA was one of Sawtooth Software's first products. The first version of ACA offered comparatively few options. Our main thought in designing it was to maximize the likelihood of useful results, which meant minimizing the number of ways users could go wrong. I think we were generally successful in that. ACA had the benefit of being developed over a period of several years, during which its predecessors were refined in dozens of actual commercial projects. Although there were some “ad hoc” aspects of the software, I think it is fair to say that “it worked.”

During the last 20 years I've had many helpful interactions with Paul Green and his colleagues. One of the most useful was a *JMR* article by Green, Kreiger, and Agarwall with suggestions about how to combine data from the self-explicated and paired comparison sections. Those suggestions led to a major revision of the product which provided additional user options.

ACA has also benefited from helpful contributions of other friendly academics, especially Greg Allenby and Peter Lenk. The ACA/HB module uses Bayesian methods to produce estimates of individual part-worths that are considerably better than the usual estimates provided by ACA.

In particular, HB provides a superior way to integrate information from the two parts of the interview. That consists of doing standard Bayesian regression where the paired comparison answers are the only data, and where the self-explicated data are used only as constraints.

I believe ACA users who are content with the ordinary utilities provided by ACA are too satisfied with their results. The results from using the HB module are enough better than those of standard ACA that I think the HB module should almost always be used.

I have been involved in one way or another with ACA for more than 30 years. During that time it has evolved from an interesting innovation to a popular tool used world-wide, and has been accepted by many organizations as a "gold standard." As I enter retirement, others are carrying on the tradition, and I believe you will see continuing developments to ACA that will further improve its usefulness.