Watching Customers Decide

Process measures add insights to choice modeling.

By Eric J. Johnson, Robert J. Meyer, Bruce G.S. Hardie, and Paul Anderson

The use of multi-attribute models to predict consumer response to new products has been one of the major success stories of marketing science. Although the family of procedures for estimating multi-attribute models is quite large—and includes both conjoint analysis and choice experiments—all are based on a simple but powerful idea: By observing the reactions of consumers to a set of hypothetical products with varying features, we can obtain quantitative estimates of the relative influence of these features on consumer evaluation of products. These estimates, in turn, can then be used to construct models that predict likely response to differing new product designs.

Although multi-attribute choice and judgment models have been used to support the product design decisions of firms in a wide variety of industries, their application nevertheless comes with a couple of important caveats. First, some argue that most modeling approaches are limited in their ability to provide detailed insights into how consumers are actually processing product attribute information when making choices. To illustrate, consumers often speak of making choices using simplifying rules that appear to use only a subset of the product attribute information that is available. For example, "I would only buy a brand name I know," or "I am looking for a computer under $1,000."

Such policies—called noncompensatory rules—turn out to be far more the norm than the exception as a description of how consumers actually make choices between sets of options. Even though traditional multi-attribute models can mimic the outcomes of noncompensatory rules, they have a much harder time diagnosing their actual structure. Such diagnosis requires information about how consumers sequentially gather, edit, and evaluate information—data that lies beyond the domain of traditional choice analyses.

The second caveat is that current methods do little to mitigate the natural hesitation by most managers to make high-stakes decisions on the basis of a single data source. While it is possible to examine the internal validity of estimated multi-attribute models by examining their ability to predict evaluations of a holdout sample of alternatives, in some cases, such validation might be considered insufficient. In the same way that firms commonly consider both quantitative and qualitative research in making decisions, managers might, in the spirit of methods like Assessor, place greater confidence in models whose parameters are arrived at by triangulation of convergent data sources.

One methodology—process-assisted choice modeling (PACMod)—addresses both these concerns. First, by incorporating observations about how information is gathered that is recorded automatically by computer in the course of the decision, we hope to add richness to the data used to estimate multi-attribute models. Second, these data allow us to understand and describe how consumers are making choices and modify the model to represent the choice process better.

Monitoring Decision Processes

Since the early to mid-'70s, there has been significant evidence that search of alternatives is incomplete. It is often conducted by attribute (initially at least) and focuses on a small number of alternatives as the choice processes extends in time. This has been interpreted as evidence for the use of multiple choice rules, particularly when working with a large number of alternatives. At first, choice is marked by use of an attribute base rule, such as "eliminations by aspects," and followed by rules such as "additive differences" or a compensatory rule that focus on the remaining two to three alternatives.

Imagine consumers walking down a supermarket aisle about to buy a breakfast cereal. As a market researcher, what would you want to know to better understand their choice processes? Most likely, you'd want to see how much attention they paid to the decision as a whole and to each alternative, using the amount and duration of search. You also might want to know something about the pattern of search: Did they compare brands on price or spend all their time looking at a single, presumably favorite brand? And finally, you might want to know something about the selectivity of the search: How many brands and attributes did they consider, and how did this change over time?

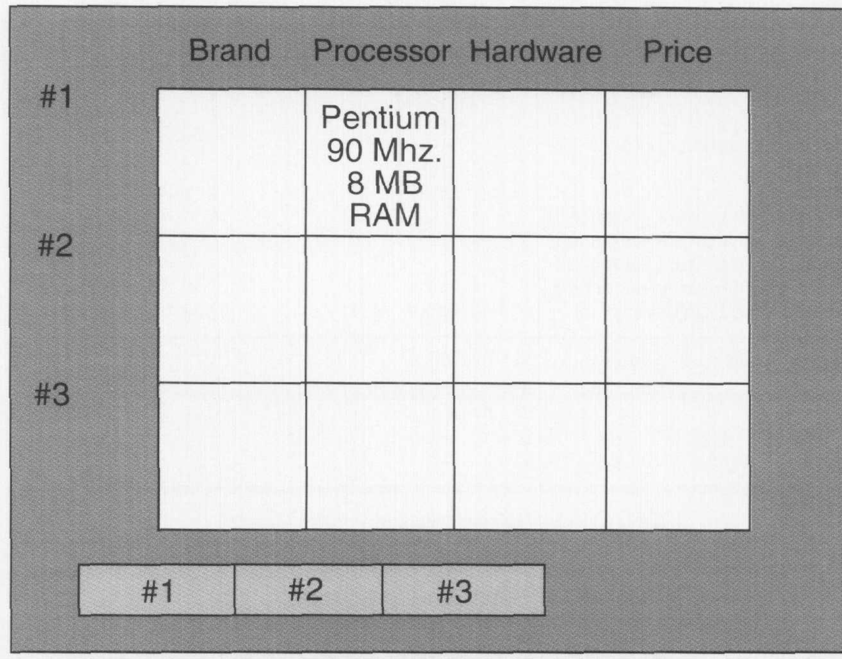
To capture this kind of data while consumers make choices, we present a computer screen in which each piece of information is accessed by moving a cursor into the box and clicking the right button on a mouse (see Exhibit 1) To translate this into the equivalent of watching the consumer shop, we have developed several indices based on experience with a similar computer system called MouseLab. We measure attention by counting the number of times a box was opened (acquisitions) and the total time that each box was opened (time). We measure the pattern of search by looking at how it varied across alternatives and calculating the standard deviation of search (search variability). To make this more concrete, this number would be small if the consumer looked equally at all the breakfast cereals in an aisle, and large if he or she spent a lot of time on one and very little on the others.

The other two statistics describing pattern are a bit more technical. The "pattern index" ranges from +1 if search is done within brands (the rows in Exhibit 1) and -1 if it is done entirely within attributes (the columns). This is calculated as: (holistic-dimensional)/(holistic+dimensional).

Finally, we capture some information about timing by examining if an alternative is examined early on in the process or survives until the end by looking at the "percent of acquisitions" (percent) made during each half of the decision. If a breakfast cereal were among the last considered, it would have a large percentage of its acquisitions during the last half of the choice; if it were eliminated early on, this number would be small.

Exhibit 1

Computer-based choice display



AN APPLICATION

Acme, a computer manufacturer, was interested in estimating how important branded ingredients—such as the "Intel Inside" logo—were in influencing consumers' selection of a computer. (Although the firm's name is disguised, the data are real.) The firm was interested in obtaining improved measurement of the effect of branded ingredients on consumer computer choices, and how these choices are made in general. To address this, a sample of 107 consumers were asked to make discrete choices from a series of 16 choice sets, each containing a differing number of hypothetical computers. Respondents participated individually in the study using portable computers in small groups of 8-10 respondents.

These computers were drawn from an orthogonal array that varied four factors:

- Brand name/price (A, B, C, D).
- Whether the computer had a branded ingredient.
- Hardware (two levels of processor speed and memory).

Exhibit 2

Choice model results, Acme Computer

	Main effects	Model w/interactions
Brand A	0.26	0.53
Brand B	-0.03	-0.19
Brand C	-0.58	-0.38
Brand D	0.38	0.04
Branded ingredient present	-0.16	-0.10[NS]
Hardware	0.80	1.11
Multimedia	0.27	0.23
Brand A w/hardware enhanced		-0.51*
Brand B w/hardware enhanced		0.26
Brand C w/hardware enhanced		0.64
Brand D w/hardware enhanced		-0.39
Log likelihood, adjusted ρ^2	-1582.08, .414	-1500.57, .437

All significant at the .05 level, except * which $p > .10$, and ns = > .10.

Exhibit 3

Process measures by attribute

Attribute	Acquisitions (number)	Time (seconds)	Variability in search	Pattern	Percent
Brand[1]	7.63[sb]	4.96[c]	18.86[a]	-.36[c]	36.7[a]
Hardware	7.27[b]	8.64[a]	14.60[c]	.04[b]	36.6[a]
Multimedia	7.47[a]	7.30[b]	15.66[bc]	.06[b]	52.9[b]
Warranty	5.60[c]	3.10[d]	15.79[b]	21[a]	58.6[c]

Means with unique superscripts within a column are significantly different, $p < .05$, by t-test using the error term from the univariate ANOVA.

- Whether the computer had multimedia enhancements.

Two sets of eight orthogonal computer profiles then were used to construct 16 choice sets of varying size, using design methods described by Jordan Louviere and George Woodworth. Unlike typical applications, however, the choice sets were presented in a brand-by-attribute matrix, as in Exhibit 1, where respondents used a computer mouse to discover the attribute values of the alternatives. Specifically, attribute information was presented either as text or, in the case of brand name and branded ingredient, as a picture. The computer then recorded the time of opening and closing of each cell along with the final choice.

To estimate the effect of the various product features upon choice, we initially subjected respondents' choices to a multinomial logit analysis, modeling choice as a linear combination of the effects of brand, and the product features of hardware, presence or absence of the branded ingredient, level of hardware, and the presence or absence of multimedia (see Exhibit 2). The coefficients, shown in the second column, shows strong brand effects along with significant positive effects of hardware level and multimedia, and a smaller negative effect of the presence of the branded ingredient. This last result suggests that the branded ingredient plays a limited role in choice and probably would not be a driver. However, because this represented an important strategic variable for Acme, managers desired a better understanding.

Process Diagnostics

Describing attributes: The results of the choice model tells us about the importance of the various attributes. It does not tell us, however, why different attributes have different levels of importance. To meet our goal of providing Acme management with more information about the choice process, we examined the process measures based upon the choices. We first looked at the process measures for each attribute, displayed in Exhibit 3.[1]

It appears that respondents did significant search in this task, but that the nature of that search varied across the attributes. For example, within these screens, the typical decision involved about 28 different acquisitions. This table tells us that different attributes are used for different things. Yet the number and type of acquisitions and their pattern differ across the different attributes. The major difference is that brand serves a different role than the other attributes. It has a negative pattern index, a lower time of search, and greater variability in search. Finally, brand name is not looked at much during the last half of the choice process; only 37% of the acquisitions occur then.

All these measures support the idea that brand is used in screening the choice set. Similarly, the hardware attribute presents some evidence that it is used early in the search, but it is not searched primarily by attribute. In contrast, warranty seems to have a very different role from brand.[2] It is looked at less frequently and for very little time. We would conclude that it is relatively unimportant to the choice process. Overall, this analysis provides significant insight for management about the role of brand name in the decision process: It is used to eliminate alternatives.

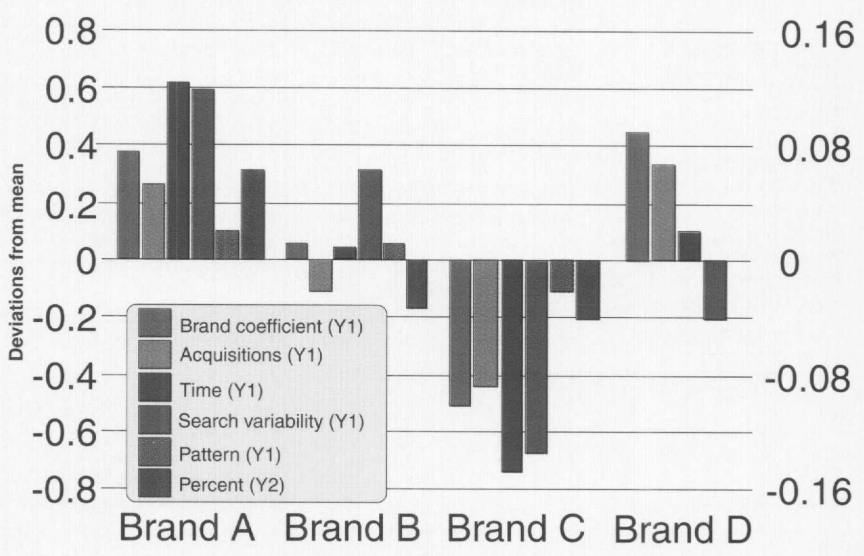
Diagnosing brands: Because brand name seems so important, it might be worthwhile to examine its role in choice further. Even though our analysis suggests that some of the brands face elimination, it does not tell us which ones. To look at this more carefully, we examined how the process variables were affected by the various brand names. For each brand, we calculated the process variables that had been looked at to date. One way of thinking about this analysis is to consider this a measure of how the brand name changes the process. By doing so, we can see if any brand systematically affects the choice process.

As Exhibit 4 indicates, there are clear winners and losers. When a computer bears the brand A name, it consistently receives more attention and is less likely to be eliminated. For the C brand, it appears that the mere presence of the name is sufficient to cause elimination. An interesting case in point is given by brand D. Although the overall attractiveness of the brand is high, as noted by the coefficients, it seems to have a different impact on the process: the name increases attention, but there is still a significant possibility that the brand will be eliminated because of other product attributes, as indicated by decreases in search variability and pattern index.

This analysis could be performed for any attribute, illustrating how the process indicators can lead to a deeper level of insight into the role these attributes—and, more important, these attribute levels—play in choice. Specifically, any choice model-based analysis of the valuation of a brand's equities could well benefit from these diagnostics.

Exhibit 4

Effect of brand name on choice process



Improving Choice Description

One interesting conclusion of this analysis of the process data is that consumers are taking shortcuts: Brand name, in particular, seems to be used to eliminate alternatives. Prior research suggests that a choice model can be easily modified to handle this if one knows which attribute is being used to eliminate alternatives. The solution is to include an interaction between the elimination attribute and the others. Given the sparse nature of discrete choice data, the Acme design could estimate only one interaction between the brand and the remaining attributes, and we chose the next most important attribute, "enhanced hardware," as indicated by both the choice model and process data.

The results, shown in Exhibit 2, not only show a significant increase in fit ($p < .00001$), but also offer an important change in interpretation: Enhanced hardware does not help, and even hurts brand A's popularity while having no significant impact on brand B. In contrast, brand C benefits greatly from the enhancements. The implication for the management of brand A, Acme, is clear: Keep it simple. For brand C, this enhanced hardware is its only chance for competing, yet it does not, on average, make up for the advantage conferred by the presence of the brand A name.

Process Segmentation

An obvious shortcoming of aggregate analyses like the ones for Acme is that they ignore differences between market segments. This is a consequence of adopting discrete choice model specifications, particularly when respondents make a limited number of choices. Although several methods—such as latent class analysis, or the use of Gibbs sampling—can be used to overcome this limitation, the process data suggest a new approach: clustering individuals with like choice processes into segments.

To do this, we conducted a cluster analysis on the process measures, and a clear three-cluster solution emerged that led to two major conclusions (see Exhibit 5 on page 40). First, there were substantial process and demographic differences. For example, the third group took almost twice as long to make decisions, had different initial preferences (favoring mail-order brands), and comprised more educated respondents. People in this group also were more experienced computer shoppers and less reliant upon salespeople for information. Second, we found we could improve fit by incorporating these differences. By estimating the main effects model

Exhibit 5

Preference, process, and demographic differences by segment

Group	Overall n = 107	Cluster 1, n = 63	Cluster 2, n = 31	Cluster 3, n = 13
Attribute preferences	Loves brands A, B, and D, hates C	Loves brands A, B, but adores D is indifferent to C	Adores brand A, likes brands B and D, but hates C	Adores brand D, indifferent to A, but hates B and C
	Hardware important	Hardware is very important	Hardware matters	Hardware matters
	Branded ingredient somewhat negative	Not at all important	Not important	Slightly negative
	Care much less about multimedia	Multimedia insignificant	Multimedia as important as hardware	Multimedia as important as hardware
Log likelihood	-1562.08	-915.66	-445.49	-176.05
Adjusted ρ^2	.414	.416	.420	.45
Process differences		Work fairly hard (29 acquisitions, 22 seconds looking time), search by attribute (pattern = -.31)	Work slightly longer (33 acquisitions, 32 seconds), but search by brand (pattern = .62) with less Search Variability	Work much longer (53 acquisitions, 49 seconds), but search by attribute (attern = -.39)
Demographic differences	Education	35% college grads	32% college grads	75% college grads
	Use salesperson as information source	62%	58%	15%
	Consider Packard Bell	11%	35%	8%
	Consider mail order brands (Dell, Gateway)	24%	9%	58%
	Experience (4-point scale), 4 = experienced	2.8	2.2	3.16

in each of the resulting clusters, we improved the fit of the data significantly, and each cluster yielded interpretable differences in coefficients.

Increasing Confidence

This case study illustrates how the process of multi-attribute modeling can be improved through the parallel collection of process-tracing data. In some ways, the results can be seen as reassuring to firms that invest heavily in the use of choice experiments to guide product design decisions; once models are properly specified, choice experiments can provide insights into how consumers make choices that converge with those yielded by process tracing methods.

But therein lies the quandary: Whether or not a model is properly specified often is difficult to discern by looking at the results of model analysis alone. Most methods used for multi-attribute modeling do not support statistically efficient exploratory searches among functional forms, leaving the analyst unsure about whether an unexpected result is a "real" one or simply reflects misspecification. By supplementing traditional choice data with process tracing measures, the analyst is a step closer to overcoming this concern by gathering independent insights into attribute effects with one data collection effort.

But the value of PACMod lies in more than its ability to serve as a reassuring guide in formal model analysis. In many applications, it will be the process-tracing data rather than the model estimates that will be seen by managers as holding the most valuable information. Specifically, rather than simply informing firms of which factors are the most important, PACMod is able to say a bit about *why*, in terms of patterns of information gathering. We saw that the effect of respondents encountering a leading name brand was not simply to pay more attention to it, but also to alter the way subsequent attribute information was gathered—shifting from a tendency to examine attribute values between brands to looking, in a confirmatory fashion, at the other attributes of that same brand.

The analysis we have reported was just one part of a major research effort on the part of Acme to better understand the role of the branded ingredient in choice. The analysis strongly indicated that there was little impact of the branded ingredient on the choice of Acme's products. As a result of this and other research, Acme accepted advertising support provided by the supplier of the branded ingredient.

Although not motivated by the research, Acme replaced the branded ingredient in some lines with a second source at a considerable savings in production costs, resulting in headlines in the trade press such "Acme rocks industry with use of second source." While the final outcome of this decision involved maintaining long-term relationships with suppliers, the discovery of flawed supplies of the branded ingredient, and an eventual recall by the supplier, we feel that the multiple sources of information about choice processes increased Acme's confidence in its decisions.

Our next efforts in applying PACMod concern the use of diagnostics in new technology products to assess the stability of the assessed preferences, specifically examining whether the preferences that we observe in choice experiments are constructed for the task or will predict preferences in the real decision environment. And another application currently in development is using these techniques to provide better understanding of choices made in computer-mediated environments. ■

ENDNOTES

[1]To confirm these apparent differences, we conducted a multivariate analysis of variance on the six measures, using as an independent variables the various attributes representing the columns of the computer display and a variable indicating respondent. Overall, the multivariate effect of attribute was significant ($F(18, 13905) = 112.58$, using Pillai's trace, $p < .0001$). In addition, each of the univariate models was significant as was the univariate F for attribute for each the process variables (all $ps < .0001$). We then conducted contrasts among the means, which are reflected differences in the way the attributes were processed for each measure and which confirm our description of the process variables.

[2]Although not part of the factorial experimental design, Acme management was interested in the role of warranty in choice. It was not included in the logistic choice models, but the process data strongly suggest that warranty plays little role in computer selection.

[3]To do this, we conducted a MANOVA with the alternative-based version of each of the five process variables as dependent variables, and with the experimental design used in the choice model as independent variables, along with a subjects factor. We then examined both the overall effect of each of the product features on the set of process measures and examined each measure more closely using univariate F tests and comparisons on the means.

ACKNOWLEDGMENT

This research was supported by the Wharton Research Forum on Electronic Commerce.

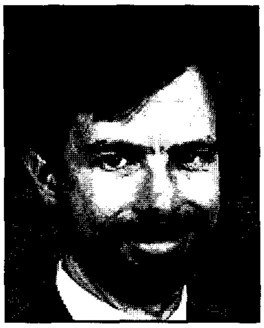
ADDITIONAL READING

Bettman, J.R., Johnson, E.J., and Payne, J. (1991). "Consumer Decision Making," in *Handbook of Consumer Behavior*, T.S. Robertson and H.H. Kassarjian eds. Englewood Cliffs: Prentice Hall.

Johnson, Eric, Meyer, Robert J., and Ghose, Sanjay (1989). "When Choice Models Fail: Compensatory Models of Noncompensatory Choice Processes," *Journal of Marketing Research*, 26 (3), 255-70

Louviere, Jordan J., and George Woodworth (1983). "Design and Analysis of Simulated Consumer Choice of Allocation Experiments: A Method Based on Aggregate Data," *Journal of Marketing Research*, 20 (November), 350-67.

Silk, A.J., and Urban G.L. (1978), "Pre-Test Market Evaluation of New Packaged Goods: A Model and Measurement Methodology," *Journal of Marketing Research*, 15 (May), 171-91.



Eric J. Johnson is a Professor of Marketing at the University of Pennsylvania's Wharton School, Philadelphia.

Robert J. Meyer is a Professor of Marketing at the University of Pennsylvania's Wharton School, Philadelphia.

Bruce G.S. Hardie is with the London School of Business.

Paul Anderson is with Yarnell Associates Inc.