Glen L. Urban & John R. Hauser

"Listening In" to Find and Explore New Combinations of Customer Needs

By "listening in" to ongoing dialogues between customers and Web-based virtual advisers (e.g., Kelley Blue Book's Auto Choice Advisor), the authors identify new product opportunities based on new combinations of customer needs. The data are available at little incremental cost and provide the scale necessary for complex products (e.g., 148 trucks and 129 customer needs in the authors' application). The authors describe and evaluate the methodologies with formal analysis, Monte Carlo simulation (calibrated on real data), and a "proof-of-concept" application in the pickup-truck category (more than 1000 Web-based respondents). The application identified opportunities for new truck platforms worth approximately $2.4 billion–$3.2 billion and $1 billion–$2 billion, respectively.

> Identifying new platform opportunities is one of the most important roles of market intelligence. Monitoring [Web-based advisers] provides a rich source of observed in-market customer behavior that complements our current inquiry tools that, by their nature, are forced to ask customers either to state their intentions before they are actually in the market or to remember after the purchase what they did (and why) when shopping for a vehicle. No form of inquiry is perfect, however; whatever its limitations, the currency [of Web-based advisers] presents a valuable source of market understanding that is already streaming by and is of great value when used appropriately.
> —Vince Barabba, General Manager of Corporate Strategy and Knowledge Development, General Motors

The advent of the Internet has given customers more information about products in diverse industries such as travel, health, automobiles, computers, home entertainment, and financial services. For example, the percentage of people using the Internet for information and advice is high in travel (70%), health (56%), and automobiles (62%). The monitoring of Internet searches, undertaken by potential customers in their own vested interests, has the potential to reveal new opportunities for new products and product platforms. In this article, we explore a set of methodologies to use this information to identify new product opportunities. Although our application is drawn from the automotive industry, the basic concepts are applicable to complex products in both consumer and business-to-business markets, such as high-end copiers, home entertainment centers, and financial services (Ulrich and Eppinger 1995).

Automobiles and trucks are indeed complex products. The investment for a new automotive platform can require as much as $1 billion–$2 billion and 1200 person-years of investment. Such investments are justified by the scale of the market. For example, with approximately 150 brands of truck on the market, the average truck needs less than 1% of the marketplace to be profitable; each share point is worth $800 million in annual revenue.

Most automotive platforms are redesigns to provide known combinations of customer benefits (i.e., needs). However, long-term survival requires that new opportunities be identified. For example, in the late 1980s, through a combination of qualitative focus groups and quantitative perceptual mapping studies, a new opportunity was identified for luxury vehicles that could haul moderate loads. Today, the luxury sport-utility-vehicle segment is one of the most profitable automotive segments. Another new product example came from leading-edge users. In the 1960s, teenagers and young adults were customizing inexpensive vintage Fords with V8 engines. Ford recognized the opportunity for inexpensive, sporty cars with large engines. The first production car in this "pony" segment, the 1964½ Mustang, sold 420,000 units in the first year ($10 billion in today's prices; ClassicPonyCars.com 2002). The 1983 Chrysler minivans are another example. Growing families needed a vehicle that could carry a 4′ × 8′ sheet of plywood, fit easily in their garages, drive like a passenger car, have a side door for small children, and incorporate a sedanlike liftgate for shopping. Chrysler sold 210,000 units in the first year and dominated the new segment for years to come (Allpar.com 2003). These are but some of the many automotive examples in which profitable new platforms filled previously unrecognized (by the auto industry) combinations of consumer needs. The firms that first identified the new combinations

---

Glen L. Urban is David Austin Professor of Marketing (e-mail: glurban@mit.edu), and John R. Hauser is Kirin Professor of Marketing (e-mail: jhauser@mit.edu), Sloan School of Management, Massachusetts Institute of Technology. This research was supported by the Sloan School of Management, the eBusiness Center, the Center for Innovation in Product Development at MIT, and the General Motors Corporation. The authors gratefully acknowledge the contributions of their industrial collaborators, research assistants, and faculty colleagues: Vince Barabba, Iakov Bart, Ahmed Benabadji, Rupa Bhagwat, Brian Bower, Brian Chan, Hann-Ching Chao, Mitul Chatterjee, Shyn-Ren Chen, Thomas Cheng, Stanley Cheung, Frank Days, Ken Dabrowski, Benson Fu, Salman Khan, Christopher Mann, Rami Musa, Joseph Kim, Ken Lynch, Bill Qualls, James Ryan, Bilal Shirazi, Jonathon Shoemaker, Fareena Sultan, Andy Tian, Xingheng Wang, and Irene Wilson. This article has benefited from presentations at the Marketing Science conferences in Wiesbaden, Germany, and Edmonton, Canada; the Marketing Science conference "New Methods for New Products"; the MIT Marketing Workshop; the New England Marketing Conference; and the Stanford Marketing Workshop. Related information is available at http://mitsloan.mit.edu/vc and http://ebusiness.mit.edu. The authors provided external funds for the use of color in the figures.

Journal of Marketing
Vol. 68 (April 2004), 72–87

of customer needs were able to exploit the opportunities profitably for many years.

Identification of new combinations of customer needs for complex products is no small challenge. For example, trucks fulfill between 100 and 150 distinct customer needs, and even more if sound and other subsystems are included. Because of the sheer magnitude of combinatorial combinations (e.g., 10^{52} in our application), existing products fulfill a tiny fraction of the potential combinations. Complex products require large samples. For example, even if we had hypotheses about a new combination of customer needs, we might still need detailed information on almost 500 or more respondents to be comfortable that a needs-combination segment is worth further investigation. Because multiple needs define a segment, it is not unusual for sample sizes in the automotive industry to approach 10,000 for targeted research and 100,000 for general searches. General Motors (GM) alone spends tens of millions of dollars each year searching for new needs combinations and studying needs combinations when they have been identified. Some studies are in the cost range of $500,000 to $1 million. Automotive firms desire methodologies that are more cost effective and that can be run continuously to identify new needs-combination opportunities as soon as they occur.

In this article, we propose methodologies that provide a practical means to find combinations of customer needs that represent profitable new opportunities. The methodologies exploit new data (i.e., clickstreams from virtual advisers) that are available at little incremental cost but provide the scale (both number of products and number of needs) that is necessary to find opportunities in complex-product categories. For example, there is a virtual adviser sponsored by GM, J.D. Power, Kelley Blue Book, and Car Talk and partly based on the methodologies in this article that has approximately 500,000 annual visitors.

We obtained the new data by “listening in” to ongoing dialogues created when customers use the Internet to search for information and advice about automotive purchases. The data are incentive compatible: Customers are seeking advice and have an incentive to reveal their needs. The virtual advisers generating the data are updated often to include new products and new customer benefits (needs), providing evolving data with which to identify new combinations of needs as soon as customers express them. We focus on the truck market to illustrate the methods. The methodologies extend readily to other complex-product categories, such as travel, medical, and office equipment.

We listen in by combining multiple stages: a Bayesian virtual adviser to obtain the data, an opportunity trigger to identify when existing trucks do not fulfill desired combinations of needs, a virtual engineer to explore and clarify the identified opportunity, a design palette to explore how customers would design their own trucks, and a clustering procedure to estimate the (rough) size of the segment of customers who desire the new combinations of needs. In this article, we illustrate each stage, examine internal validity with Monte Carlo analyses, and provide an example based on a sample of more than 1000 respondents. This “proof-of-concept” research was performed parallel to existing methods, yet it identified a key segment at a much lower cost. It also implied the existence of a segment, still being explored, that existing methods may have missed. We begin by discussing how listening in complements existing methods.

Existing Methods to Identify Profitable Combinations of Customer Needs

Because so much is at stake, strategic marketing and marketing research groups invest heavily in identifying new opportunities. They speak to leading-edge users, maintain and monitor user groups, sponsor special racing events, monitor chat rooms and user groups, and use various qualitative and ethnographic methods (Barabba 2004; Barabba and Zaltman 1991; Griffin and Hauser 1993; Gutman 1992). For example, automotive firms invest heavily in quantitative methods such as conjoint analyses; activities, interests, and opinions (AIO) studies; and large-scale “clinics” in which customers view and react to prototypes and concepts (Green and Srinivasan 1990; Plummer 1974; Urban, Weinberg, and Hauser 1996). Table 1 summarizes characteristics of existing methods and listening in. The cost and sample-size data are typical for the automotive industry; they are based on our experience and discussions with auto executives and consultants.[1]

The methods in Table 1 are complementary. For example, qualitative and ethnography interviews are powerful methods to probe in-depth once the research is focused, but they are an expensive means to search for combinations of needs that might be desired by less than 1% of the market. Conjoint analyses provide accurate estimates of the importance of customer needs, but they are most effective when they are targeted to approximately 10 to 20 needs. Even adaptive methods cannot handle all the needs that describe a truck. Furthermore, AIO studies are designed to examine the entire market for new combinations of needs, but they are expensive, performed infrequently, and tend not to collect data on gaps in customer needs. In contrast, AIO studies provide critical input to virtual advisers. Truck clinics provide the most realistic stimuli to customers. They are designed carefully to forecast sales before launch, but their primary use is confirmatory rather than exploratory.

Listening in fills a gap in existing methods by making it feasible to use inexpensive and readily available data to search large numbers of customer needs to find combinations of customer needs that are desired but not currently fulfilled by existing trucks. More important, unlike AIO studies, listening in can immediately and automatically target both quantitative and qualitative questions to explore further the new combinations of customer needs. Because listening in runs continuously and is updated periodically with new vehicles and benefits (needs), it provides an early warning of new needs-combination segments as soon as they appear in the market.

Tailored interviewing (TI) has characteristics that are similar to the Bayesian virtual adviser. Both TI and the vir-

[1]Table 1 also includes tailored interviewing, an approach that shows promise for automotive applications, especially for the segmentation gearbox used in the virtual adviser.

TABLE 1
Complementary Methods for Understanding Customer-Needs Combinations: Truck Example

	Data Source	Incremental Cost per Study	Number of Features or Stated Needs per Study	Number of Vehicles per Study	Number of Feature Combinations	Includes In-Depth Probes
Qualitative and ethnographic interviews[a]	5–10 groups of 5–10 customers per segment	$40,000–$50,000	50–100	5–10	Open-ended	Yes
Tailored Interviews[b]						
Segmentation	800 personal interviews	$80,000	73 scales	7 segments	—	—
Interest or intent	512 (calibration) and 235 mail questionnaires	$10,000–$15,000	12–15 items (Phase 2)	Single scale	—	—
AIO studies[a]	100,000 mailed questionnaires	$500,000	114	150	—	—
Conjoint analyses[a]	300 online or in-person interviews	$50,000–$100,000	10–20	5–10	10^6	—
Truck clinics[a]	300 central-facility personal interviews	$500,000	40–50	10–20	—	Yes
Listening in[c]						
Bayesian adviser	Track online	$10,000–$20,000	36	148	—	—
Opportunity trigger	Track online	Included	—	—	10^{15}	—
Virtual engineer	Invite online	Included	79	—	10^{31}	Yes
Design palette	Invite online	Included	14	—	10^6	—
Clustering	Track online	Included	—	—	—	—

[a]Data are as cited in article and/or typical for the automotive industry. We rounded some estimates for confidentiality.
[b]Data are from the work of Kamakura and Wedel (1995) and Singh, Howell, and Rhoads (1990). Automotive cost estimates are based on sample sizes in *Journal of Marketing Research* articles. Cost per respondent is typical for the industry, as is estimated by auto industry executives and consultants.
[c]Experience is based on the pilot study. The numbers of vehicles, needs, and combinations may increase in subsequent applications.

tual adviser classify respondents (e.g., into seven segments, as in the work of Kamakura and Wedel [1995]; into three most preferred trucks [of 148] in our application). There are other technical differences that we discuss in the next section. A key conceptual difference is that to be practical in the truck market, the virtual adviser must be updated almost continuously as new trucks enter the market or as new features are added to the question banks. Although both methods assign respondents with posterior probabilities, the virtual adviser relies on Bayesian methods to update probabilities and uses data from multiple sources, whereas TI relies on a calibration survey and uses maximum-likelihood methods (Kamakura and Wedel 1995, Equations 3–7). Each method works well in its target application.

Listening in is not a panacea, nor can it operate without complementary methods. For example, although the virtual engineer contains qualitative probes, subsequent qualitative and ethnographic research provides greater depth on a segment when it has been identified. Similarly, when new needs combinations have been uncovered, conjoint analyses search the combinations in greater detail and quantify the importance of the alternative needs. Although listening in provides first-order forecasts, truck clinics provide the accuracy necessary before \$1 billion–\$2 billion is committed to a project. We illustrate in a stylized way how listening in complements existing methods for two practical situations in truck markets. In practice, applications are more iterative and include other methods (Urban and Hauser 1993).

Identify opportunities for a new truck platform:
Listening in ⇒ qualitative interviews ⇒ conjoint analysis ⇒ truck clinics ⇒ launch.

Monitor marketplace changes for vehicle "refresh" opportunities:
Listening in ⇒ conjoint analysis ⇒ truck clinics ⇒ launch.

Tapping Data from Virtual Advisers (Web-Based Searches)

Virtual-adviser data are extensive, available at little incremental cost, and underused as a means to identify unfulfilled combinations of customer needs. Web sites such as Kelley Blue Book (http://www.kbb.com), Microsoft Autos (http://autos.msn.com), Edmund's (http://www.edmunds.com), Autobytel (http://www.autobytel.com), Autoweb (http://www.autoweb.com), NADA (http://www.nadaguides.com), and Vehix (http://www.vehix.com) have changed the way that customers search for information on cars and trucks. Of all new-vehicle buyers, 62% search online before buying a vehicle (J.D. Power and Associates 2001). This search rate has increased from 54% in 2000 and from 40% in 1999. The most important and most accessed Internet content is information about vehicle options and features. Notably, although customers prefer independent sites for pricing and general evaluation, they prefer manufacturers' sites, by more than a two-to-one margin, for feature and option information (J.D. Power and Associates 2001, p. E16).

Virtual advisers come in many varieties, including comparators, which array choice alternatives by features (Epinions.com); feature-specifiers, which ask consumers for preferred levels of features and search the database for products that meet the feature specifications (Kelly Blue Book's online recommendation tool); configurators with detailed specifications and costs for the chosen set of detailed product features (http://configurator.carprices.com/autoadvisors); collaborative filters, which recommend products based on correlations of previous purchases by similar customers (Amazon.com); and utility maximizers, which use methods similar to conjoint analysis to weight features (Activebuyersguide.com). Other advisers use real people who consumers can access by e-mail (Mayohealth.org) or in live chat rooms (Nordstom.com).

The listening-in methodology relies on data from a Bayesian virtual adviser, which is a method that is well-matched to the opportunity trigger mechanism. However, the virtual engineer, the design palette, and the clustering are not limited to working with a Bayesian virtual adviser. These methodologies can work with any virtual adviser that provides recommendations at any point in the questioning sequence and that links customers' responses to benefits that the customers derive from vehicles.

A Bayesian Virtual Adviser

The Bayesian virtual adviser was developed as a prototype for a major automotive manufacturer; a commercial system based, in part, on this adviser is now in place on the Web. This virtual adviser combines two methods to recommend a set of four vehicles to customers: a segmentation gearbox and a Bayesian adviser. The segmentation gearbox divides people into segments on the basis of grouping and assignment rules.[2] The grouping is based on a cluster analysis of a 114-item AIO questionnaire sent to 100,000 respondents (76 personal viewpoints and 38 preferred vehicle characteristics, including styling and design). The automotive manufacturer's AIO study identified 48 segments, of which 25 were relevant to pickup trucks. Customers were assigned to segments on the basis of answers about their desires for features and options such as comfort, passenger capacity, and prestige as well as about their anticipated use of the truck. In the virtual adviser, one of the four recommended vehicles was the vehicle bought most often by the segment to which the customer was assigned. However, because the segmentation gearbox is designed to allocate people to segments rather than identify new opportunities, it is not the focus of this article. Instead, we focus on the Bayesian adviser that recommends three of the four vehicles.

Bayesian Adviser

The basic concepts behind the Bayesian adviser are (1) to select sets of questions, known as question banks, such that the answers provide the most information about which vehicle to recommend and (2) to update the probabilities that describe the likelihoods that each vehicle will be most pre-

---

[2]The industry term "gearbox" is an analogy. Just as the gearbox in a car matches engine speed to wheel speed, the segmentation questions match the manufacturer's vehicles to the customer.

ferred by the customer after each question bank.[3] Figure 1, Panel A, illustrates the opening screen of the virtual adviser (a neighbor who is a contractor and who has bought many trucks over the years), and Figure 1, Panel B, illustrates one of the question banks asked of customers. We describe the Bayesian updating mechanism and then describe how it can be used to select the maximum-information question bank. We subsequently indicate how we obtained both the conditional and the prior probabilities.

We begin with the notation. We let Q be a set of question banks indexed from q = 1 to N. For each question bank, q, r_q indexes the potential responses to that question bank, where r_q is a nominal variable with values from 1 to n_q. If there is more than one question in a question bank, n_q represents the number of possible combinations of answers. If one of the questions includes a continuous sliding scale, it is discretized to a finite number of categories.

For each customer, the order of the question banks is chosen adaptively. For a given customer, R_{q-1} is the set of question banks up to but not including question bank q. The variable v_j indicates vehicles from 1 to V. At any point in the adviser's questioning sequence, we are interested in the likelihood that the customer will prefer vehicle j after having been asked question bank q. We indicate this likelihood by $P(v_j|R_{q-1}, r_q)$.

Suppose that from previous surveys, we have available the conditional probabilities of how customers, who prefer each vehicle, will answer the question banks. We then can use Bayes' theorem to update recommendations.[4]

$$(1)\qquad P(v_j \mid R_{q-1}, r_q) = \frac{P(r_q \mid v_j, R_{q-1})\, P(v_j \mid R_{q-1})}{\sum_{j=1}^{V} P(r_q \mid v_j, R_{q-1})\, P(v_j \mid R_{q-1})},$$

where $P(v_j|R_{q-1})$ is the virtual adviser's recommendation probability to the customer for vehicle v_j before asking the qth question bank.

However, even with data from full-scale surveys, such as an AIO questionnaire with 100,000 responses, use of Equation 1 is not feasible because the number of potential combinations of responses grows exponentially with the number of question banks. For example, in our study, the dimensionality of R_N, the number of unique paths through the adviser's questions, is 1.4×10^{15}. Fortunately, we can make Equation 1 feasible based on the property of local independence. This property appears reasonable for our data and has proved robust in simulations and applications in the TI literature (e.g., Kamakura and Wedel 1995, Equation 11; Singh, Howell, and Rhoades 1990, Equation 8). Local independence recognizes that there are nonzero correlations across vehicles in the answers to the question banks; customers who prefer a full-sized truck may also prefer a diesel engine. Indeed, it is this combination of preferences on which the adviser bases its recommendations. However, if we limit ourselves to customers who prefer a Ford F350 Supercab, for those customers, responses to the "size" question bank are approximately statistically independent of the responses to the "engine type" question bank. This enables us to write $P(r_q, R_{q-1}| v_j) \cong P(r_q|v_j)\, P(r_{q-1}|v_j) \ldots P(r_1|v_j)$, which implies that $P(r_q|v_j) \cong P(r_q|v_j, R_{q-1})$ by the laws of conditional probability. Using this property, we rewrite Equation 1, in which we recursively obtain $P(v_j|R_{q-1})$, as follows:

$$(2)\qquad P(v_j \mid R_{q-1}, r_q) \cong \frac{P(r_q \mid v_j)\, P(v_j \mid R_{q-1})}{\sum_{j=1}^{V} P(r_q \mid v_j)\, P(v_j \mid R_{q-1})}.$$

Figure 2 gives a simplified example for one customer of the evolution of the recommendation probability. The current recommendation is on the left-hand side, and the probability that the customer will purchase that recommended vehicle is on the right-hand side. Also listed on the left-hand side are the question bank and parts of the answer. For example, after the second question bank on engine size, the customer answers "four cylinders." If the customer were to stop answering question banks and request a recommendation, the adviser would recommend the Mazda B2300 and forecast a .0735 probability that the customer would purchase the Mazda B2300. In Figure 2, the probability of purchase increases for the most preferred truck after each question bank is answered. Note that the recommended vehicle changes after the fifth question bank and again after the eighth question bank.

Question Bank Selection

To select the next question bank, the virtual adviser attempts to gain as much information as possible from the customer. For example, if after reviewing the responses, the adviser decides that a question bank on towing capacity is likely to make one truck more highly probable and all other trucks less probable, that question bank might be a good candidate to ask next. To do this, we turn to formal theory in which information is defined as the logarithm of the relative odds (e.g., Gallagher 1968). That is, the information, $I(v_j|r_q, R_{q-1})$, provided by the response to question bank q equals $\log [P(v_j|R_{q-1}, r_q)/P(v_j|R_{q-1})]$. This definition has several nice theoretical properties, including that (1) under an equal proportional loss rule, information always increases when the probability of the maximum-choice truck increases; (2) the expected information is maximized for the true probabilities; and (3) the information measure rewards systems that provide more finely grained estimates (Kullback 1954; Savage 1971).[5]

To compute the expected information, we take the expectation over all possible responses to question bank q and over all possible vehicles. The information that we expect from question bank q is given in Equation 3:

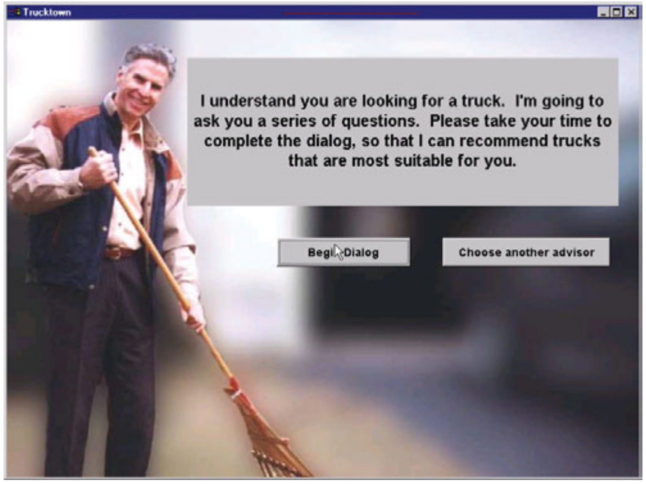
[3]The global set of question banks from which the algorithm selects is drawn from cluster analyses of the ongoing AIO surveys, supplemented with managerial judgment. The set of question banks evolves on the basis of ongoing market intelligence. These methods are state of the art but standard marketing research practice. They are not the focus of this article.

[4]In most equations, we suppress the individual customer subscript, i, for simplicity.

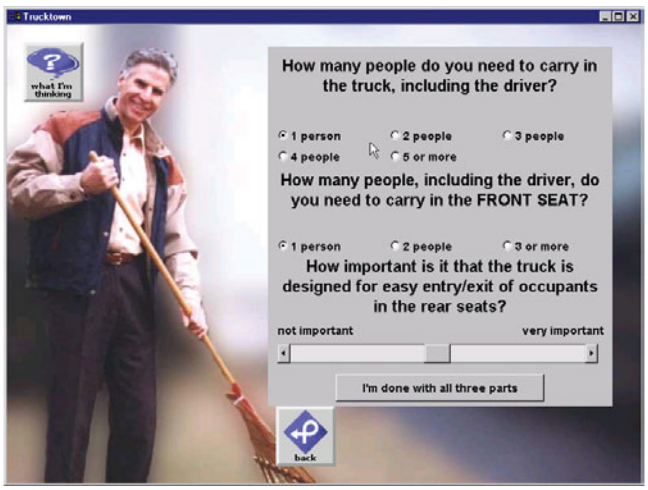
[5]For applications in marketing of reward functions based on information theory, see Hauser (1978) and Herniter (1973). For applications in psychology, see Prelec (2001).

FIGURE 1
Example Question Banks Asked by Bayesian Virtual Adviser

A: Introductory Screen



B: Example Question Bank



$$
\text{(3)} \quad EI(q \mid R_{q-1}) = \sum_{j=1}^{V} P(v_j \mid R_{q-1}) \sum_{r_q=1}^{n_q} P(r_q \mid v_j, R_{q-1}) \times \log \frac{P(v_j \mid r_q, R_{q-1})}{P(v_j \mid R_{q-1})}.
$$

We use a two-step look-ahead algorithm. For each potential question bank and response on Step 1, the adviser computes the best second question bank and the expected information for that question bank. It then selects the Step 1 question bank with the highest contingent expected information.

Initial Calibration

Two estimates are necessary and sufficient for the virtual adviser: prior probabilities, $P(v_j)$, and conditional response probabilities, $P(r_q|v_j)$. The virtual adviser obtains the prior probabilities for each individual from a logit model based on five truck characteristics: price, fuel economy, performance, reliability, and safety. Each customer is asked initial constant-sum, self-explicated importance weights (w_c) for these characteristics. (The prior weights are obtained from questions that are asked before the question banks illustrated in Figure 2.) We estimated the prior probabilities with Equation 4, where w_c is the importance for the cth characteristic, x_{jc} is the value of characteristic c for vehicle v_j, and β is a scaling parameter:

$$
\text{(4)} \quad P(v_j) = \frac{e^{\beta \sum_{c=1}^{5} w_c x_{jc}}}{\sum_{j=1}^{V} e^{\beta \sum_{c=1}^{5} w_c x_{jc}}}.
$$

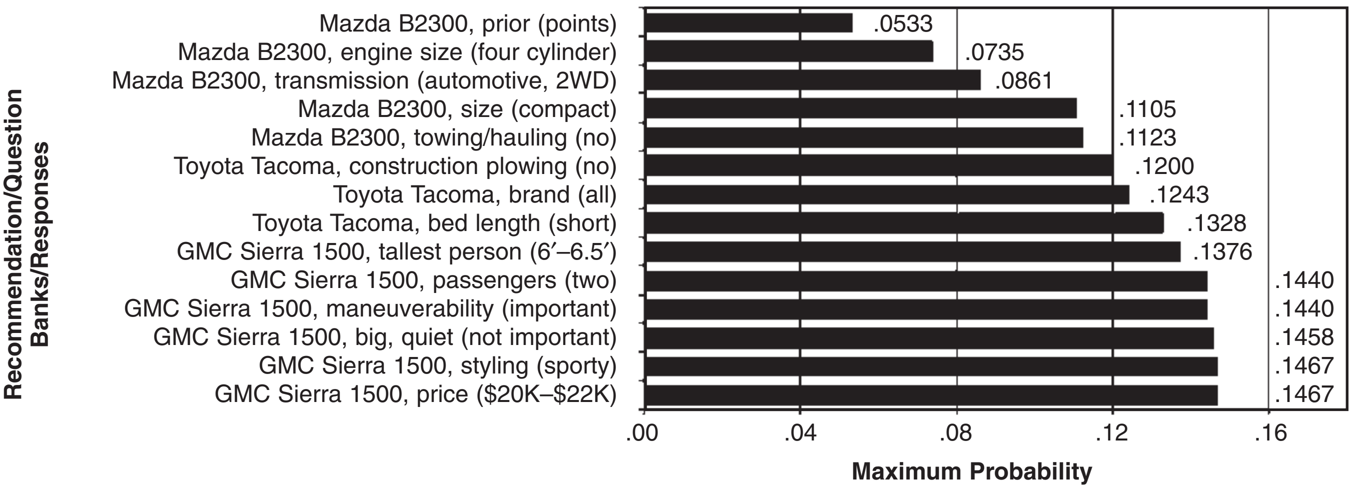
We obtained the characteristic values for each existing vehicle and the scaling parameters from archival data and managers' and engineers' judgments. For example, prior surveys of owners help establish that the Toyota Tacoma 4 × 4 (regular cab) has a rating of 1.087 on fuel economy and a rating of 1.241 on performance. For the GMC Sonoma two-wheel-drive regular cab, the corresponding ratings are 2.116 and .525, respectively (data are disguised slightly). We synthesized the actual data from "an ongoing global effort" by the manufacturer "to understand consumers' needs and wants related to motor vehicles" (quotes from a proprietary study). Part of this ongoing global effort included data from the AIO questionnaire described previously (76 personal viewpoints and 38 vehicle characteristics). When new vehicles become available, managers and engineers provide temporary estimates of the x_{jc}'s.

The conditional response probabilities are based on the ongoing AIO surveys, supplemented when necessary by experienced managers and engineers. For example, the survey data suggest that customers who prefer the Toyota Tacoma 4 × 4 (regular cab) are likely to answer that they prefer a four-wheel-drive vehicle 84% of the time. They are likely to answer that they prefer two-wheel drive only 16% of the time. Table 2 illustrates data, disguised slightly, on conditional probabilities for numbers of passengers that are obtained from AIO studies. The data, $P(r_q|v_j)$, are sufficient for the updating equations (Equations 2 and 3) if they are available for all question banks in the virtual adviser.

Evolving Question Banks

Virtual advisers and listening-in are not one-shot studies. Markets evolve as customer needs change and as technology improves. Each year brings changing features and new truck brands. To advise customers and identify new opportunities effectively, it must be relatively simple to update the prior and conditional probabilities with data from multiple sources. For example, suppose that four-wheel steering becomes a feature that is important to customers (and a feature that helps the adviser recommend a truck). Suppose further that some truck brands begin offering this feature for the 2003 model year. We add a question bank on steering to the set of available trucks. Because of the local independence property, we need obtain only incremental data for the new question banks. We need to know how owners of each

FIGURE 2
Evolution of Updated Recommendation Probabilities After Question Banks



Notes: Abbreviated consumer responses to question bank are in parentheses. 2WD = two-wheel drive.

TABLE 2
Conditional Probabilities Obtained from AIO Surveys and Supplemented with Judgment

	Conditional Probability $P(r_q|v_j)$ (%)				
Number of Passengers	**Chevy Avalanche 2WD**	**Chevy Silverado 2500 2WD**	**GMC Sonoma 4WD Crew Cab**	**(148 Vehicles)**	**Dodge Ram Club 4WD**
1	5%	25%	15%	…	10%
2	15	25	5	…	15
3	25	25	15	…	25
4	25	15	25	…	25
5–6	30	10	25	…	25

Notes: Data are disguised. 2WD = two-wheel drive; 4WD = four-wheel drive.

truck brand will rate *their* vehicles on the new question bank. For new truck brands, we need to know how owners of the new brands will rate their vehicles on the characteristic values (x_{jc}) and how they will answer each question bank, $P(r_q|v_j)$. We obtained the data from the periodic AIO surveys and from other sources, such as one-time surveys and judgment. In essence, the virtual adviser (and listening in) free rides on surveys undertaken by the manufacturer for other purposes. This adaptability is a key feature that is necessary for practical application and represents a conceptual difference between the Bayesian virtual adviser and TI. The former uses Bayesian methods to incorporate new data from multiple sources, whereas the latter relies on maximum-likelihood estimates obtained in a calibration survey. Each method is matched to its application domain. However, further research might combine these relative strengths into an improved methodology.

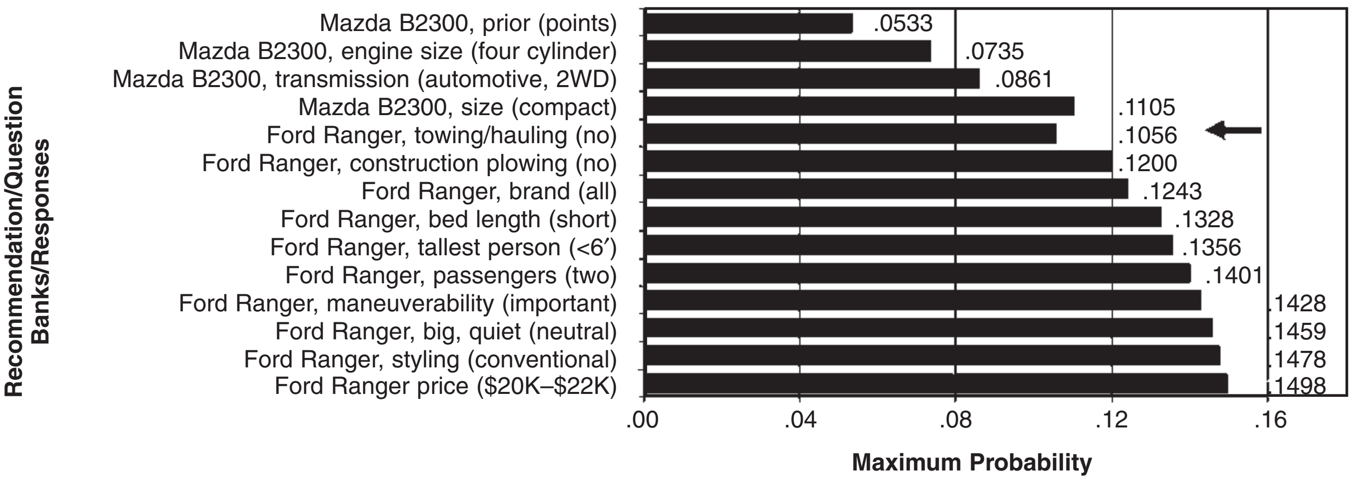
Opportunity Trigger Mechanism

The next stages of listening in identify when opportunities exist and identify the combinations of customer needs that are not satisfied by existing vehicles.

Trigger Mechanism to Identify When Opportunities Exist

For many customers, an existing vehicle will fulfill their needs, and the updated recommendation probabilities will evolve smoothly as in Figure 2. Existing vehicles satisfy the needs combinations these customers desire. However, for some customers, their answers to question banks reveal inconsistencies. For example, suppose that (1) the customer has already answered constant-sum importance question banks, which indicate that reliability and low price are important (price 30 points, performance 10 points, fuel economy 20 points, reliability 30 points, and safety 10 points), and (2) the customer's subsequent answers suggest an interest in a small truck with a four-cylinder engine, two-wheel drive, and automatic transmission. Through the first four question banks, the Mazda B2300 fits these preferences best (see Figure 3, first four bars from top). Given these answers, the virtual adviser decides that further information on towing and hauling will clarify recommendations. The adviser expects that the customer will want to haul or tow relatively light loads, such as small garden equipment or a Jet Ski. Knowing the exact towing and hauling needs will

FIGURE 3
Example Use of the Opportunity Trigger



Notes: Abbreviated consumer responses to question bank are in parentheses. 2WD = two-wheel drive.

help the adviser decide among several otherwise comparable light-duty trucks.

However, suppose that the customer says that he or she plans to use the truck to haul heavy materials and to tow a large motorboat (weighing 6500 pounds). No existing light-duty truck can tow such heavy loads effectively and safely. In contrast, no truck that can tow such heavy loads can fill the customer's requirements as expressed in previous question banks. If enough customers desire these combinations of features, this may be an opportunity worth investigating: a light-duty truck that can occasionally haul heavy materials or tow heavy loads. Note that the goal is to define the opportunity by needs (light duty, haul heavy materials) rather than features (V8 engine). In this way, new vehicles can satisfy the newly identified combinations of customer needs with features that may or may not be available in existing vehicles.

The intuition in this example is that the question bank on towing and hauling revealed something about the customer's underlying needs. This new information suggests that the customer is not satisfied with the needs combinations provided by existing trucks; the virtual adviser will need to revise its best-truck recommendation probability downward. This drop in the maximum recommendation probability becomes a trigger for further investigation. We illustrate this trigger mechanism with an arrow in the dialogue in Figure 3. The fifth question bank, which included questions about towing and hauling, causes the most preferred vehicle to change from the Mazda to a Ford Ranger (a slightly larger, more powerful compact truck). Utility drops because this more powerful compact truck is an insufficient compromise to meet both the towing and hauling requirements and the requirements expressed in the first four question banks (it has a six-cylinder engine and is more expensive). A full-sized truck, such as the Chevrolet Silverado 1500, can fulfill the towing and hauling requirements, but the adviser does not recommend the Silverado because it has poor ratings on the other desired features. After further question banks, the recommendation probabilities in Figure 3 again increase because the Ford Ranger fulfills the additional requirements.

The intuitive idea in Figure 3 has appeal, but before we incorporate the trigger mechanism, we must investigate it further. For example, the posterior probability might drop because there is error in the customer's response. If the trigger mechanism is too sensitive, it might identify many false need-conflicts, and the true need-conflicts might be lost in the noise. In contrast, if it is not sensitive enough, the trigger mechanism might miss opportunities. We show subsequently, through simulation, how to select a sensitivity level for the trigger mechanism such that segments of customers desiring known combinations of needs are recovered with sufficient precision. In the simulations, we begin with real data for the conditional probabilities and create known segments. We then add error and examine how various sensitivity levels balance false positives and false negatives. The simulations demonstrate that calibration is feasible and that the performance of the listening-in mechanism is reasonably robust in the face of response errors. It is also reasonably robust with respect to the sensitivity levels chosen for the trigger mechanism. Having thus established a reasonable degree of internal validity, we are more confident in applying the methodology to real data.

The other issue is theoretical. The intuition assumes that a drop in posterior probability identifies a conflict in the desired customer needs that are fulfilled by existing vehicles. If a question bank affected only the vehicle that was recommended before the qth question bank and if that same vehicle were recommended after the qth question bank, then most random utility models would suggest that a probability drop was an indicator of an underlying utility drop. For example, both the logit and the probit models have this property. However, each question bank can affect the probabilities of all 148 vehicles and change the identity of the recommended vehicle on the basis of the qth question bank. We demonstrate formally in the Appendix that the intuition still holds. If the qth question bank does not change the identity of the recommended vehicle, a drop in posterior probability is a necessary and sufficient condition indicating that the recommended vehicle has characteristics in conflict with the customer's preferences. The more complex issue is when the qth question bank changes the identity of the recommended vehicle. We show formally that if the recommended vehicle changes and the posterior probability drops, it must be the case that a truck with mixed characteristics would have higher utility than the truck recommended either before or after the qth question bank. We also show that the mixed-characteristic truck that is better for the customer is not an existing truck.

Analyses to Identify Which Combinations of Customer Needs Are Not Satisfied

When a probability drop identifies a potential conflict, we seek further information to identify which customer needs are in conflict. We consider a null hypothesis that the existing trucks satisfy (almost all) customer-needs combinations. This hypothesis implies that if two truck characteristics are positively correlated among existing trucks, we expect them to be positively correlated among customers' preferences, as revealed by their answers to the questions banks. For example, on the basis of existing trucks, we expect that there is a positive correlation across vehicles of the probabilities that a customer will (1) use the truck for towing heavy loads and (2) prefer a rugged body style for that vehicle. In addition, we expect that there is a negative correlation of the probabilities that a customer will (1) use the truck for towing heavy loads and (2) prefer a compact body style. Because no existing truck satisfies these needs simultaneously, recommendation probabilities will drop when the customer requests a compact truck that can tow heavy loads (see the Appendix).

This means that we can identify the needs combinations that caused the drop by examining negative correlations among expected answers to the question banks for the questions answered by customers who experienced a probability drop. The probability drop challenges the null hypothesis and its implications. That is, customers who experience a probability drop want some combinations of customer needs that are negatively correlated in the existing market. To find the desired combinations from the set of all negatively correlated combinations, we limit our search to the need combinations evaluated by customers with probability drops.

Formally, $\rho_{r_q r_p}$ is the correlation across vehicles of the conditional probabilities of a customer answering r_q to question bank q and answering r_p to question bank p,[6] and P is the matrix of these correlations (here P is a capital ρ). Whenever a probability drop implies a potential opportunity, the listening-in algorithm examines all correlations corresponding to that customer's answers to the first q question banks $(R_{q-1} \cup r_q)$ and flags the ones that are highly negative (less than –.30 in our application). Such negative correlations indicate why the (triggered) customer's desired benefits (needs) are not fulfilled by existing trucks (subject to statistical confidence). The level of the flagging mechanism is set with simulation.

The opportunity trigger identifies the customers who have combinations of needs that are not satisfied, and it flags specific entries in the P matrix to identify combinations of needs that represent new opportunities. The combinations of needs are a working hypothesis for a new opportunity. However, before the automotive firm can act on the working hypothesis, it needs further information about the potential opportunity, because the number of questions the virtual adviser uses is, by necessity, a compromise between efficient recommendation (fewer questions) and probes for new needs combinations (more questions). To understand and explore the opportunity more completely, listening in complements the virtual adviser and the trigger mechanism.

A Virtual Engineer Clarifies the Opportunity

The virtual engineer (VE) concentrates its questions to obtain relevant, more-detailed information about combinations of customer needs. The VE asks relatively few questions of each targeted customer (six screens in our application), but across many customers, its questions span the needs space. In our application, the VE explores an additional 79 features beyond the 36 features explored in the virtual adviser. As is the virtual adviser, the VE is designed to be flexible; its questions are updated continuously without the need to recommission large-scale AIO surveys.

The concept of a VE is simple; its implementation difficult. To be useful, the VE must ask the customer questions that inform the engineering design decisions that are necessary to design a truck to meet the customers' newly identified (potential) combination of needs. To be credible to the customer, the VE must ask questions in a nontechnical manner that pertains to how the customer uses the truck. Naturally, the VE evolves through application, but we describe here the process by which the initial VE questions are created.

An engineering design team from a major automotive manufacturer considered the basic engineering problem imposed by potential conflicting needs. The team then generated the questions that it would need answered to clarify the opportunity and to decide among basic solutions to conflicts. The engineering team members formulated the questions that they would ask the customer if they were participating in the dialogue between the adviser and customer. For example, if a customer wants a compact truck that can tow a large boat, the engineering team would ask about the type of boat (e.g., modest sailboat, large motorboat, multiple Jet Skis) and the weight of the boat that the customer plans to tow. The engineering team would also ask the customer why he or she wants a compact truck (e.g., low price, tight parking, high maneuverability, fuel economy). All engineering questions are then rephrased into "customer language."

In addition to the questions identified by the engineering team, the VE includes open-ended dialogues that enable the customer to elaborate further the reasons underlying the previously unidentified combinations of needs. Figure 4 illus-

[6]Such correlations *across* vehicles are consistent with local independence, which assumes response independence conditioned on a given vehicle. Local independence enables customers to be heterogeneous across vehicles in their answers to the question banks.

trates a sample dialogue in which the VE introduces himself, asks about a conflict, gathers quantitative data, and asks for open-ended comments. In this example, the conflict is between a full-sized truck and a six-cylinder engine.

A Design Palette Solicits Customer Solutions to Potential Conflicts

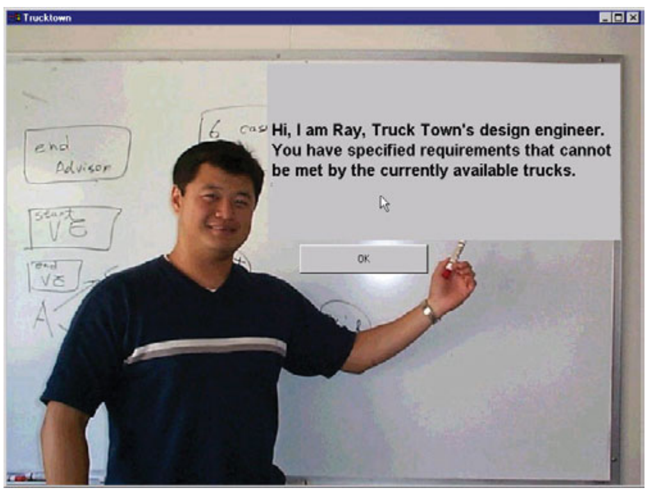
We supplement the VE with a design palette (DP) that covers 14 features. The DP's perspective is the customer's own solutions (von Hippel 1986). The DP is similar to innovation toolkits, configurators, and choice boards that enable customers to mix and match features (Dahan and Hauser 2002; Hauser and Toubia 2003; Liechty, Ramaswamy, and Cohen 2001; von Hippel 2001).

The DP is illustrated in Figure 5. The customer (1) receives instructions, (2) changes the size of the truck, and (3) changes the color. For brevity, we do not show the many intermediate steps, some of which include new state-of-the-art truck features, such as four-wheel steering and extrawide frames. However, changes are not free for the customer. There are sophisticated engineering and cost models underlying the DP. For example, if the customer changes the size of the truck, the price, fuel economy, and towing/payload capacity change accordingly. After completing the redesign, the customer is given the opportunity to indicate whether and by how much he or she prefers the new design. (The customer may not prefer the new design because of accumulated sticker shock or because of a holistic judgment of the final truck.) In the empirical application that we describe subsequently, 73% of the respondents who completed the exercise indicated that they would purchase their custom-designed truck were it available.[7]

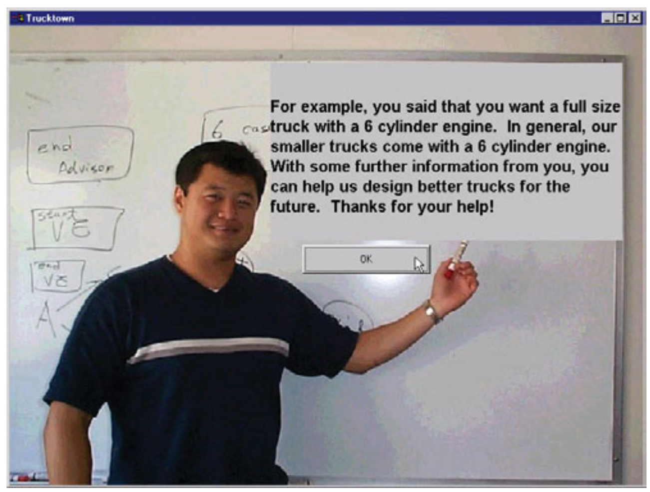
[7]Becuase of self-preference learning, memory accessibility, and context effects, the preference for the self-designed truck may be inflated (Simmons, Bickart, and Lynch 1993; Tourangeau, Rips, and Rasinski 2000). This does not diminish the value of the DP as a means to clarify opportunities.

FIGURE 4
Virtual Engineer

A: Introductory Screen



B: Example Dialogue



C: Specific Questions to Elaborate

D: Open-Ended Question

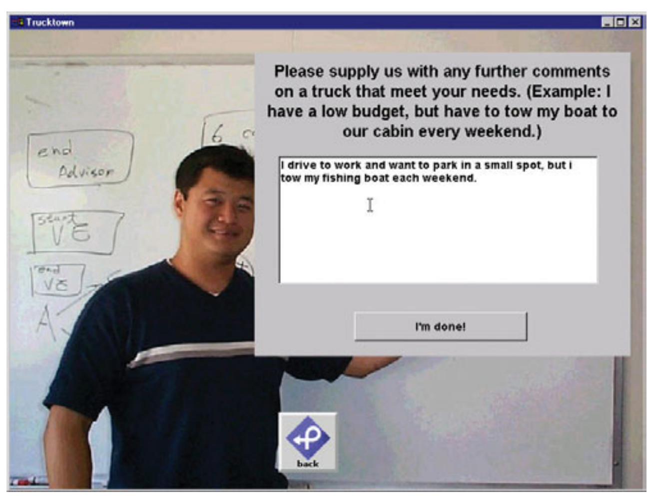
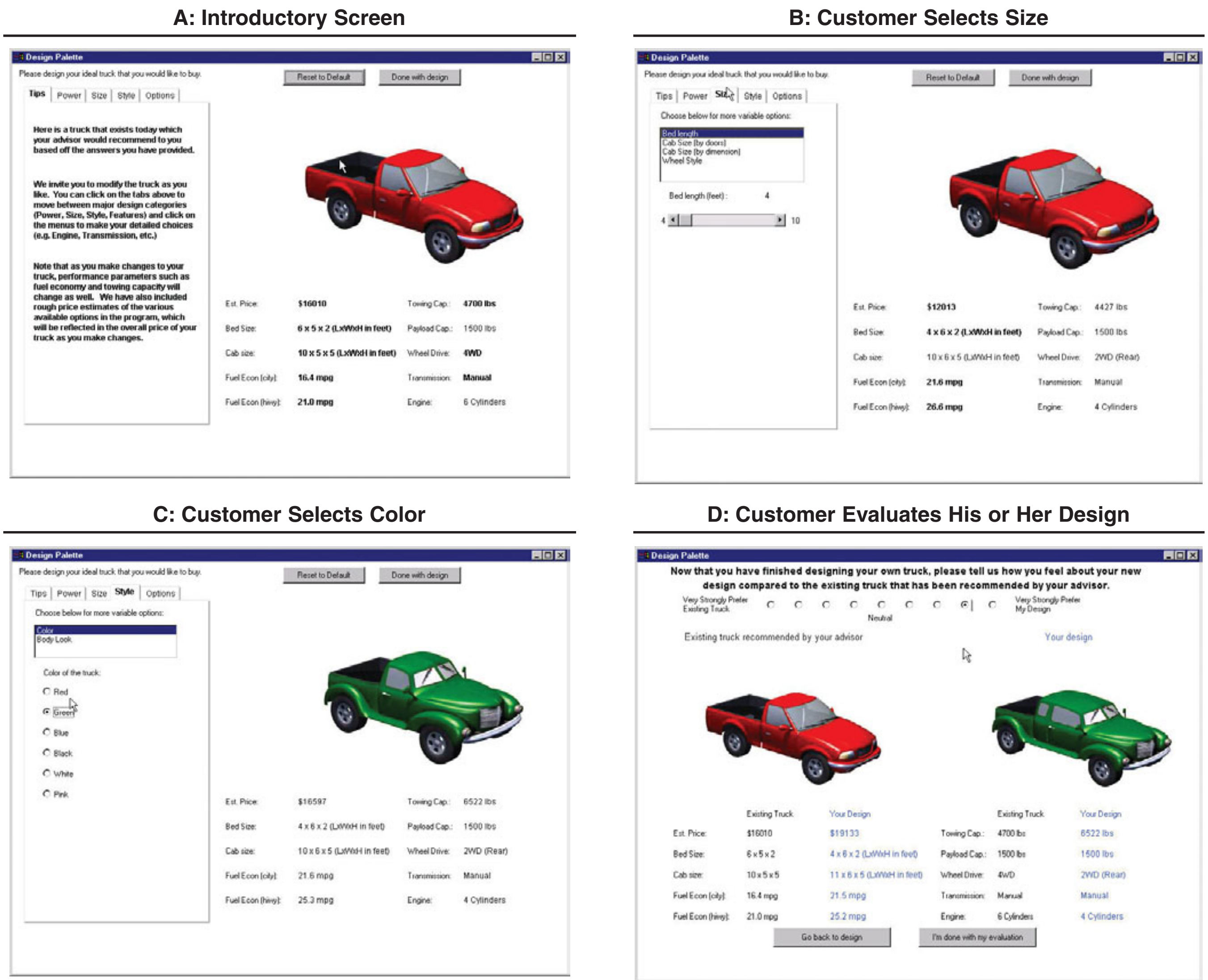


FIGURE 5
Design Palette



In general, DPs are evolving rapidly. For example, there is a system that enables the customer to adjust the length of the hood of a car or truck while the software automatically ensures the integrity of other design elements, such as the windshield angle and window shape. The customer simply clicks on the hood and drags it forward or clicks on the front bumper and pushes it back. Using this advanced DP, the customer easily creates a "Euro" sports design (short front overhang, high truck deck, low overall height) that is pleasing to the eye and incorporates many design heuristics. In contrast, by lengthening the front overhang and the hood the customer creates a classic look with a long sloping back to the truck. The software is sufficiently advanced that the customer can then rotate the model in all directions for a full three-dimensional view.

Together, the virtual adviser, VE, and DP explore 129 customer needs (10^{52} combinations, many of which are multilevel). The detailed data help the firm understand the customer-need conflicts that led some customers to experience a probability drop. The philosophy behind this listening-in search differs from conjoint analysis. Conjoint analysis collects data on the importance of customer needs and searches to find needs combinations that satisfy a minimum share of the market profitably. In contrast, listening in monitors needs requests to identify when customers request combinations of needs that are not fulfilled by existing trucks. After the opportunities are identified, they can be explored further with conjoint analysis.

Initial Sizing of the Opportunity

The next stage of listening in groups customers according to their unmet combinations of needs as revealed through flagging components of the P matrix (supplemented with the VE and DP for interpretation). This estimate of market potential is a rough indicator, but it is sufficient to identify potential opportunities for the fuzzy front end of an iterative product development process. The firm evaluates the opportunities further with targeted qualitative and quantitative research.

Suppose that A_i represents customer i's answers to the question banks. For each A_i, we identify a subset, P_i, of the P matrix that represents strongly negative correlations. By clustering *triggered* respondents on P_i, we identify groups of customers with similar combinations of desired needs that are not fulfilled (on average) by existing trucks. Subject to the caveat of self-selected customers, the size of the cluster as a fraction of the initial sample is a rough indicator of the size of the segment that desires the identified combinations of needs.[8] To simulate a new truck design, we define a concept truck by the needs it fulfills as reflected by customers' answers to the question banks, $P(r_q|v_j)$. These data are sufficient to calculate revised posterior probabilities for all trucks, including the new-truck concept (Equation 2). Averaging of the revised posterior probabilities over respondents provides a rough estimate of the potential market share for the new concept truck.

Monte Carlo Simulations: Sensitivity to Error and the Trigger Mechanism

If successful, listening in will affect billion-dollar decisions on new truck platforms. Before we can be confident in its application, we must address the following issues: First, we want to know whether listening in can recover opportunities from noisy data. This issue is best addressed with simulation because we can specify known segments of customers who have unmet needs combinations. Second, applications require that the opportunity trigger be calibrated. Here, too, simulation is best to determine the best trigger sensitivity. Relevance and external validity are better addressed with a proof-of-concept application in which we listen in to real customers in a pilot study to determine whether unmet combinations of needs can be identified. We hope that the pilot study at least can identify combinations of needs that were discovered in parallel by other studies (at much greater expense). Recall that truck manufacturers routinely spend tens of millions of dollars annually on market research.

Simulation Methodology

We use the conditional probabilities, $P(r_q|v_j)$, and P-matrix correlations based on the 100,000-respondent AIO study and supplemental managerial judgment. On the basis of the proof-of-concept study we describe subsequently, we select three segments of customers whose needs are satisfied by existing trucks (e.g., full-sized trucks that can tow and haul large loads). The three segments provide a baseline from which to test whether the methodology identifies false opportunities. Next, we generate six segments with combinations of needs that are not satisfied by existing trucks. We define their responses to the question banks to be consistent with their desired benefits (needs). We attempt to test whether listening in can recover these segments from noisy data. Because of the multiple stages of listening in, this is far from ensured. In total, we generate nine customer segments of 500 respondents each, for a total of 4500 simulated respondents.

We next add errors to the customers' responses. For the r_q's, which are nominal variables, we randomly select E% of the questions to be answered incorrectly. The incorrect answers are distributed among the remaining categories according to a uniform distribution. For the w_c's, which are interval-scaled variables (mean = 20), we simulate response error by adding a zero-mean, normally distributed response error such that the standard deviation of the error equals a specified number of points (e). For simplicity, we truncate negative self-explicated importances that, fortunately, occur with low probability. We then apply the listening-in equations to each of the 4500 simulated respondents. For clustering the P matrix, we use a k-means nontree clustering algorithm based on the Euclidean norm defined on the matrix of negative correlations from triggered respondents (respondents by potential conflict pairs; details are available on request).

Internal Validity: Testing Recovery of Unmet Needs Combinations from Noisy Data

As an initial test of internal validity, we add moderate noise where e = 5 points and E = 10%. We use a relatively sensitive opportunity trigger; we record conflict correlations whenever $P(v_1|r_q, R_{q-1}) - P(v_1|R_{q-1}) \leq .00005$. We subsequently examine sensitivity to this parameter.

Table 3 suggests that listening in can recover known needs combinations from moderately noisy data. The entries indicate the number of respondents from a true segment (rows) that were assigned to a cluster (columns). We examine Table 3 at the macro and micro levels.

The managerial focus is at the macro level. First, we notice the diagonal nature of the data in Table 3; even with noise in the data, listening in identified all five segments. Second, we examine the unmet combinations of needs that defined each segment. For example, the first known segment was defined by four need conflicts: compact truck/tow large loads, compact truck/haul large loads, four-cylinder engine/tow large loads, and a four-cylinder engine/haul large loads. In Cluster 1, the percentages of respondents who had these needs were 95.9%, 82.4%, 77.3%, and 73.3%, respectively. We identified no other need conflict for more than 9.4% of the Cluster 1 respondents. We obtained similar results for the other five known clusters. We identified no false-positive needs combinations at the macro level (Clusters 8 and 9 are redundant with Cluster 6).

At the micro level, we classified 82.7% of the respondents correctly. Most of the misclassifications were respondents who were classified falsely into the null segment because of errors in their responses. The simulation identified 21,096 conflict pairs compared with only 16,500 true conflict pairs: 14% were false negatives, and 36% were false positives. Thus, response errors affect the classification of specific respondents. Fortunately, the macro-level identification of unmet needs combinations appears robust with respect to the micro errors. We now test whether this insight

[8]There is self-selection because customers choose to initiate dialogues with virtual advisers. Nonetheless, a large fraction of self-selected customers might be an important opportunity. We expect less self-selection as more truck customers use the Web to search for information.

generalizes to other levels of errors (e and E) and other sensitivities of the opportunity trigger.

Setting the Sensitivity of the Opportunity Trigger and Its Relative Robustness

Table 4 repeats the simulations for various trigger sensitivities (t) that vary from extremely sensitive (t = .00000) to extremely insensitive (t = .10000). At both the macro and micro levels, listening in is relatively robust with respect to the trigger level for t ≤ .001. For larger sensitivities, performance degrades. For extremely high t, all opportunities are missed. On the basis of Table 4 and simulations with other levels of error, we recommend a sensitive trigger. The exact level is less critical as long as the level is less than .001.

Sensitivity to the Level of Response Errors

We now explore the sensitivity of listening in to response errors in the constant-sum question banks (e) and the nominal question banks (E). We examine performance at both the macro level (percentage of needs combinations identified) and the micro level (percentage of respondents classified correctly). Table 5 suggests that performance is relatively insensitive to errors in the priors (w_c's), even for errors that are 50% of the mean response (ten points). For a Bayesian system, we did not find this surprising; the impact of the priors diminishes as more question banks are answered. However, performance is sensitive to errors in the nominal question banks, with clear degradation at a 20% error. Such an error rate would correspond to one of five respondents saying that they want a compact truck when they actually want a large truck. Table 5 indicates that care must be taken in Web design to engage customers with clear questions so that error rates (E) remain at 10% or lower.[9]

Summary

Together, Tables 3, 4, and 5 suggest that a reasonable level of internal validity exists despite errors in both the prior preferences and the responses to the question banks. As long as the trigger level is relatively sensitive (≤.001) and the nominal error is moderate (≤10%), listening in can identify known segments of customers who desire combinations of needs that existing trucks do not meet. Recovery is not perfect when there are response errors, but this level of recovery should be sufficient for the fuzzy front end of product development, especially when final managerial decisions are refined with subsequent qualitative and quantitative data.

Proof-of-Concept Application and Test

Before bringing online listening in to a situation in which more than 350,000 customers are tracked annually, we believed it was important to test the methodology in a pilot

[9]There appears to be a slight anomaly in Table 5. For E = 20%, classification and identification appear to increase slightly with errors in the self-explicated importance. This happens because the combination of errors pushes more respondents to the no-conflict clusters. As a result, a few more no-conflict respondents are classified correctly, making it easier to achieve a majority in the remaining clusters. Neither difference is significant at the .05 level with a two-tailed t-test.

TABLE 3
Results of the Simulated Cluster Analysis

	Number of Respondents Classified to Each Cluster									
Needs Combinations	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Total**
Compact truck, large loads	**418**	0	0	1	0	0	81	0	0	500
Sporty full-sized, short bed	1	**422**	0	0	0	0	77	0	0	500
Compact truck, diesel	0	0	**401**	0	0	0	99	0	0	500
Full-sized, extrashort bed	1	0	0	**346**	0	0	153	0	0	500
Compact truck, ten cylinders	3	27	0	0	**336**	0	134	0	0	500
Full-sized, maneuverable	0	2	0	0	0	**346**	92	43	17	500
Null segment	43	0	2	1	0	0	**1454**	0	0	1500

Notes: Each known segment desired multiple needs combinations. Here, we list examples for each segment. The largest number in each row is in boldface.

TABLE 4
Calibrating the Opportunity Trigger

Trigger Level	Percentage of Respondents Classified Correctly	Percentage of Opportunities Identified Correctly	Percentage of Needs-Combinations Segments Identified	False Opportunities Identified
t = .00000	82.73	100	100	0
t = .00005	82.73	100	100	0
t = .00010	82.69	100	100	0
t = .00100	82.69	100	100	0
t = .01000	56.69	63.6	63.4	0
t = .10000	33.33	0	0	0

TABLE 5
Sensitivity to Response Errors

Response Errors (Updating)	Errors in the Self-Explicated Importance (Priors) e = 0 Points	e = 5 Points	e = 10 Points
Macro Level: Percentage of Unmet Needs Combinations Identified Correctly			
E = 0%	100%	100%	100%
E = 10%	100%	100%	100%
E = 20%	93.9%	75.8%	81.8%
Micro Level: Percentage of Respondents Classified Correctly			
E = 0%	100%	99.9%	99.9%
E = 10%	82.9%	82.7%	81.8%
E = 20%	61.6%	55.0%	56.8%

test with real customers. In August 2001, an automotive manufacturer sponsored a study in which 1092 pickup-truck customers were recruited from the Harris Interactive Panel and given a $20 incentive to participate in the test.[10] On average, each customer spent 45 minutes with the virtual adviser, DP, and VE (when triggered). Most customers found the experience worthwhile. Customers trusted the virtual adviser by an eight-to-one margin over dealers and would be more likely to purchase a vehicle recommended by the virtual adviser by a four-to-one margin over a vehicle recommended by a dealer. For the DP, 78% of participants found using it an enjoyable experience, and 82% believed it was a serious exercise. When the VE was triggered, 88% of participants found the questions easy to answer, and 77% believed that the VE related well to their needs. Notably, 56% of the participants reported that they would pay for the advice provided by the virtual adviser if it were included in the price of the pickup truck that they purchased as a result of using the adviser.

With a sensitive trigger, the most common pairwise conflicts were a maneuverable full-sized truck (38%), a compact truck that could tow and haul heavy materials (14%), and a full-sized truck with a six-cylinder engine (7%). Two segments of customers were identified that expressed unmet combinations of needs. Segment 1 requested large trucks but indicated a desire for maneuverability. Segment 1 consisted of two groups: customers who wanted a top-of-the-line truck and customers who wanted a standard full-sized pickup truck. Segment 2 requested a compact truck that could tow and haul heavy loads. Table 6 provides more detail on Segment 1. From the VE, we learned that respondents use full-sized trucks for city driving. Large trucks fulfill critical needs for large passenger capacity and large payloads. However, the respondents also desired maneuverability: combinations of benefits (needs) that are not available with existing trucks.

The DP explored Segment 1's desires further. The features that they changed most often were truck height (6′ to 7′), truck width (6′ to 7′), and steering (two-wheel to four-wheel steering). This suggests that these customers desire an even larger truck but that they would be interested in four-wheel steering to gain maneuverability. Using the methods described previously for market sizing, we estimated the potential market share of a full-sized truck with four-wheel steering. On the basis of cost models, we calculated that the extra features would increase the manufacturer's suggested retail price by $3,000. For this concept truck, the listening-in equations estimate a market-share increase for the manufacturer of 3%–4% (we coded the exact value for confidentiality).[11] Such a $2.4 billion–$3.2 billion annual opportunity is worth further investigation. In addition, a compact truck with heavy-duty hauling and towing is estimated to be a $1 billion–$2 billion opportunity (values are coded). Technically, the benefit (needs) combinations are feasible with the use of a small truck platform that has a strong frame, transmission, and engine.

After we completed our study, we learned that an automotive manufacturer was in the process of introducing four-wheel steering to improve the maneuverability of its top-of-the-line pickup truck, which was previously unknown to us. This combination of needs had been identified with traditional methods (Table 1) but at a significantly greater cost. This truck is now selling well. We plan to monitor the sales of this truck to determine whether its sales are in the rough range predicted by the market-sizing equations. We found no indication that traditional methods identified the need for a basic truck with four-wheel steering. We plan to monitor whether traditional methods confirm such a combination of needs.

[10]We based this initial test on a stratified random sample of the panel. For this test, all customers were given the opportunity to use the DP.

[11]We obtain rough forecasts by adding a full-sized maneuverable pickup truck to the choice sets of the needs-segment customers. We obtain $P(r_q|v_j)$ for the new vehicle by assuming a profile similar to an existing vehicle except for the critical responses on the size and maneuverability questions, which we changed to be consistent with the vehicle being both full-sized and maneuverable. The iterative use of Equation 1 provides the estimates.

TABLE 6
Elaboration of Customer Needs for a Full-Sized Maneuverable Pickup Truck

Why I Need a Maneuverable Pickup Truck		Why I Need a Full-Sized Pickup Truck	
Frequent city driving	66%	Large passenger capacity	73%
Tight parking	58%	Large payloads	50%
I make many U-turns	26%	Full-sized style	39%
Too many traffic jams	28%		

Summary, Discussion, and Further Research

In this article, we explore a methodology to listen in on customer dialogues with virtual advisers to identify combinations of customer needs that are not fulfilled by existing trucks. Monte Carlo analyses suggest that listening in is internally valid and relatively robust with respect to response errors and trigger sensitivity. A proof-of-concept demonstration suggests that unmet needs combinations for real respondents can be identified.

As with all methodologies, listening in will benefit from continuous improvement. Each stage can be improved; better methods to identify priors, more efficient look-ahead algorithms, improved calibration of the trigger mechanism, and better indicators of conflicting needs all can benefit from further research. The dialogues, the user interfaces, and the presentation of stimuli are all areas of potential improvement. For example, work is now underway to put more stretch into the DP and to give the virtual adviser and the VE personalities based on "talking heads." The various stages of listening in are designed to be modular. Further research might explore other advisers, triggering mechanisms, means to identify and size segments (e.g., latent structure analysis), and applications (e.g., telecommunications, consumer electronics, travel services, financial services).

Appendix
Formal Derivation of Trigger Mechanism

The listening-in methodology uses a trigger mechanism to invoke the VE and DP. We argue intuitively in the text that such a drop in the recommendation probability (Equation 1) is an indication that existing trucks do not fulfill desired combinations of customer needs. Here, we demonstrate with a formal analytical model that such a drop identifies opportunities. The issue is not trivial because a question bank, q, potentially affects the updated utilities of *each and every product* in the market, not just the recommended product. The formal analysis identifies the net effects.

Although our application uses complex question banks for 148 trucks, we can illustrate the basic principles with N = 3 and a dichotomous question bank. (Our propositions generalize to analogs for larger N and for polychotomous question banks, but the notation is cumbersome.) Following the text, j indexes the vehicles. Without loss of generality, v_1 is the recommended product after question bank q – 1. In addition, $\vec{x}_j$ represents customer benefits (needs) that are not affected by question bank q, and $\vec{y}_j$ represents customer benefits (needs) that are affected by question bank q. In this formulation, we treat price as a characteristic, and it can be in either $\vec{x}_j$ or $\vec{y}_j$ (for motivation, see Hauser and Urban 1996). Following Blackorby, Primont, and Russell (1975), we model preferences using a utility tree such that $u(\vec{x}_j, \vec{y}_j) = u_x(\vec{x}_j) + u_y(\vec{y}_j) + \varepsilon$, where ε is a Gumbel-distributed error term that represents the uncertainty in utility due to question banks that have not yet been asked (or may never need to be asked). For simplicity, we assume that trucks with $\vec{y}_j = \vec{y}_{good}$ experience an increase in utility, and trucks with $\vec{y}_j = \vec{y}_{bad}$ experience a decrease in utility. (The dichotomous question bank reveals which customer benefits are desired.) We let v_2 be a surrogate for products with desirable characteristics and v_3 be a surrogate for products with undesirable characteristics (as revealed by question bank q). Following McFadden (1974), we write the recommendation probabilities in more fundamental utility-theory terms (where V is the total number of vehicles):

$$\text{(A1)}\qquad P(v_j \mid r_q, R_{q-1}) = \frac{e^{u_x(\vec{x}_j)+u_y(\vec{y}_j)}}{\sum_{m=1}^{V} e^{u_x(\vec{x}_m)+u_y(\vec{y}_m)}}$$

$$= \frac{e^{u_x(\vec{x}_j)+u_y(\vec{y}_j)}}{e^{u_x(\vec{x}_1)+u_y(\vec{y}_1)} + e^{u_x(\vec{x}_2)+u_y(\vec{y}_{good})} + e^{u_x(\vec{x}_3)+u_y(\vec{y}_{bad})}}.$$

After question bank q, two situations can occur: The recommended truck remains v_1 or it becomes v_2. It cannot become v_3, because even if $\vec{y}_1 = \vec{y}_{bad}$, v_1 would still be preferred over v_3. The following propositions address the two situations. Together, they indicate that whenever the recommendation probability drops, an opportunity exists for a new higher-utility truck with mixed characteristics.

P_1: If the recommended truck after question bank q is the same truck as that recommended after question bank q – 1, then v_1 has undesirable characteristics ($\vec{y}_1 = \vec{y}_{bad}$) if and only if $P(v_1|r_q, R_{q-1})$ decreases. If the probability decreases, a new truck with mixed characteristics has higher utility than does the recommended truck. That new truck is not currently available in the marketplace.

P_2: If the recommended truck after question bank q is different from the truck recommended after question bank q – 1 and if the recommendation probability decreases, then v_1 has undesirable characteristics ($\vec{y}_1 = \vec{y}_{bad}$). A new truck with mixed characteristics has higher utility than both the recommended truck after q – 1 question banks and the recommended truck after q question banks. That new truck is not currently available in the marketplace.

Proofs

Straightforward algebra establishes that $P(v_1|r_q, R_{q-1}) - P(v_1|R_{q-1})$ is proportional to $e^{u_x(\vec{x}_1)+u_x(\vec{x}_3)+u_y(\vec{y}_{good})} - e^{u_x(\vec{x}_1)+u_x(\vec{x}_3)+u_y(\vec{y}_{bad})} \geq 0$ if $\vec{y}_1 = \vec{y}_{good}$ and that $e^{u_x(\vec{x}_1)+u_x(\vec{x}_2)+u_y(\vec{y}_{bad})} - e^{u_x(\vec{x}_1)+u_x(\vec{x}_2)+u_y(\vec{y}_{good})} \leq 0$ if $\vec{y}_1 = \vec{y}_{bad}$. Algebra also establishes that the proportionality (denominator) is positive. This establishes the first statement in P_1 and implies that $\vec{y}_1 = \vec{y}_{bad}$ if the probability drops. Because $u_1(\vec{x}_1) + u_y(\vec{y}_{good}) > u_1(\vec{x}_1) + u_y(\vec{y}_{bad})$, a new product with $\vec{x}_1$ and $\vec{y}_{good}$ has higher utility. If the recommended truck changes after question bank q, then $P(v_1|r_q, R_{q-1}) < P(v_2|r_q, R_{q-1})$, and because the recommendation probability decreases, we have $P(v_2|r_q, R_{q-1}) < P(v_1|R_{q-1})$. Thus, $P(v_1|r_q, R_{q-1}) < P(v_1|R_{q-1})$, and by P_1, we have $\vec{y}_1 = \vec{y}_{bad}$. This establishes the first result in P_2. Because v_1 was recommended before question bank q, we have $u_x(\vec{x}_1) > u_x(\vec{x}_2)$; by supposition, we have $u_y(\vec{y}_{good}) - u_y(\vec{y}_{bad}) > 0$. Thus, a product with the features $\vec{x}_1$ and $\vec{y}_2$ will have higher utility than either v_1 or v_2. This establishes the second result in P_2. In both propositions, we know that the new truck does not currently exist, because if it were available, it would have higher utility and thus would have been recommended.

Generalizations

If there are n_2 trucks similar to v_2 and n_3 trucks similar to v_3, the analogs to P_1 and P_2 are readily proved. The numbers n_2 and n_3 enter the equations for $P(v_1|r_q, R_{q-1}) - P(v_1|R_{q-1})$, but the basic proofs remain intact. If there are many trucks with $\vec{y}_{good}$ or $\vec{y}_{bad}$ but different $\vec{x}_j$, the expressions for $P(v_1|r_q, R_{q-1}) - P(v_1|R_{q-1})$ include more terms, but each can be proved to have the correct sign (i.e., increases if $\vec{y}_{good}$ and decreases if $\vec{y}_{bad}$). With these changes, the remaining portions of the proofs follow as we have showed.

We use Equation 1 to make the proofs transparent. Both propositions can be generalized to other probability models with the appropriate characteristics. We leave the details of these generalizations to readers.

REFERENCES

Allpar.com (2003), "A Brief History of the Minivan," (accessed December 1, 2004), [available at http://www.allpar.com/model/m/history.html].

Balasubramanian, Siva K. and Wagner A. Kamakura (1989), "Measuring Consumer Attitudes Toward the Marketplace with Tailored Interviews," *Journal of Marketing Research*, 26 (August), 311–26.

Barabba, Vincent P. (2004), *Surviving Transformation: Lessons from GM's Surprising Turnaround*. New York: Oxford University Press.

——— and Gerald Zaltman (1991), *Hearing the Voice of the Market: Competitive Advantage Through Creative Use of Information*. Boston: Harvard Business School Press.

Blackorby, C., D. Primont, and R.R. Russell (1975), "Budgeting, Decentralization, and Aggregation," *Annals of Economic Social Measurement*, 4 (1), 23–24.

ClassicPonyCars.com (2002), "History of the Ford Mustang," (accessed December 31, 2002), [available at http://www.classicponycars.com/history.html].

Dahan, Ely and John R. Hauser (2002), "The Virtual Customer," *Journal of Product Innovation Management*, 19 (September), 332–53.

Gallagher, Robert (1968), *Information Theory and Reliable Communication*. New York: John Wiley & Sons.

Green, Paul E. and V. Srinivasan (1990), "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice," *Journal of Marketing*, 54 (October), 3–19.

Griffin, Abbie J. and John R. Hauser (1993), "The Voice of the Customer," *Marketing Science*, 12 (Winter), 1–27.

Gutman, Jonathan (1982), "A Means–End Chain Model Based on Customer Categorization Processes," *Journal of Marketing*, 46 (Spring), 60–72.

Hauser, John R. (1978), "Testing the Accuracy, Usefulness, and Significance of Probabilistic Models: An Information Theoretic Approach," *Operations Research*, 26 (May–June), 406–421.

——— and Olivier Toubia (2003), "Properties of Preference Questions: Utility Balance, Choice Balance, Configurators, and M-Efficiency," working paper, Massachusetts Institute of Technology Center for Innovation in Product Development.

——— and Glen L. Urban (1986), "Value Priority Hypotheses for Consumer Budget Plans," *Journal of Consumer Research*, 12 (March), 446–62.

Herniter, Jerry (1973), "An Entropy Model of Brand Purchase Behavior," *Journal of Marketing Research*, 10 (November), 361–75.

J.D. Power and Associates (2001), *New Autoshopper.com Study*. Agoura Hills, CA: J.D. Power and Associates.

Kamakura, Wagner A. and Michel Wedel (1995), "Life-Style Segmentation with Tailored Interviewing," *Journal of Marketing Research*, 32 (August), 308–317.

Kullback, S. (1954), *Information Theory and Statistics*. New York: John Wiley & Sons.

Liechty, John, Venkatram Ramaswamy, and Steven Cohen (2001), "Choice Menus for Mass Customization: An Experimental Approach for Analyzing Customer Demand with an Application to a Web-Based Information Service," *Journal of Marketing Research*, 38 (May), 183–96.

McFadden, Daniel (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, P. Zarembka, ed. New York: Academic Press, 105–142.

Plummer, Joseph T. (1974), "The Concept and Application of Life Style Segmentation," *Journal of Marketing*, 38 (January), 33–37.

Prelec, Drazen (2001), "A Two-Person Scoring Rule for Subjective Reports," working paper, Sloan School of Management, Massachusetts Institute of Technology.

Savage, L.J. (1971), "Elicitation of Personal Probabilities and Expectations," *Journal of the American Statistical Association*, 66, 783–801.

Simmons, Carolyn J., Barbara A. Bickart, and John G. Lynch Jr. (1993), "Capturing and Creating Public Opinion in Survey Research," *Journal of Consumer Research*, 30 (September), 316–29.

Singh, Jagdip, Roy D. Howell, and Gary K. Rhoads (1990), "Adaptive Designs for Likert-Type Data: An Approach for Implementing Marketing Surveys," *Journal of Marketing Research*, 27 (August) 304–321.

Tourangeau, Roger, Lance J. Rips, and Kenneth A. Rasinski (2000), *The Psychology of Survey Response*. New York: Cambridge University Press.

Ulrich, Karl T. and Steven D. Eppinger (1995), *Product Design and Development*. New York: McGraw-Hill.

Urban, Glen L. and John R. Hauser (1993), *Design and Marketing of New Products*, 2d ed. Englewood Cliffs, NJ: Prentice Hall.

———, Bruce Weinberg, and John R. Hauser (1996), "Premarket Forecasting of Really-New Products," *Journal of Marketing*, 60, (January), 47–60.

von Hippel, Eric (1986), "Lead Users: A Source of Novel Product Concepts," *Management Science*, 32 (7), 791–805.

——— (2001), "Perspective: User Toolkits for Innovation," *Journal of Product Innovation Management*, 18 (4), 247–57.

Zaltman, Gerald (1997), "Rethinking Market Research: Putting People Back In," *Journal of Marketing Research*, 23 (November), 424–37.