MARKETING SCIENCE
Vol. 26, No. 4, July–August 2007, pp. 532–549
ISSN 0732-2399 | EISSN 1526-548X | 07 | 2604 | 0532

informs

DOI 10.1287/mksc.1060.0213

Greedoid-Based Noncompensatory Inference

Michael Yee
Lincoln Laboratory, Massachusetts Institute of Technology, 244 Wood Street,
Lexington, Massachusetts 02420-9108, myee@ll.mit.edu

Ely Dahan
UCLA Anderson School, 110 Westwood Plaza, B-514,
Los Angeles, California 90095, edahan@ucla.edu

John R. Hauser
Massachusetts Institute of Technology, E40-179,
50 Memorial Drive, Cambridge, Massachusetts 02142, jhauser@mit.edu

James Orlin
Massachusetts Institute of Technology, E53-363, 50 Memorial Drive,
Cambridge, Massachusetts 02142, jorlin@mit.edu

Greedoid languages provide a basis to infer best-fitting noncompensatory decision rules from full-rank conjoint data or partial-rank data such as consider-then-rank, consider-only, or choice data. Potential decision rules include elimination by aspects, acceptance by aspects, lexicographic by features, and a mixed-rule lexicographic by aspects (LBA) that nests the other rules. We provide a dynamic program that makes estimation practical for a moderately large numbers of aspects.

We test greedoid methods with applications to SmartPhones (339 respondents, both full-rank and consider-then-rank data) and computers (201 respondents from Lenk et al. 1996). We compare LBA to two compensatory benchmarks: hierarchical Bayes ranked logit (HBRL) and LINMAP. For each benchmark, we consider an unconstrained model and a model constrained so that aspects are truly compensatory. For both data sets, LBA predicts (new task) holdouts at least as well as compensatory methods for the majority of the respondents. LBA's relative predictive ability increases (ranks and choices) if the task is full rank rather than consider then rank. LBA's relative predictive ability does not change if (1) we allow respondents to presort profiles, or (2) we increase the number of profiles in a consider-then-rank task from 16 to 32. We examine trade-offs between effort and accuracy for the type of task and the number of profiles.

Key words: lexicography; noncompensatory decision rules; choice heuristics; optimization methods in marketing; conjoint analysis; product development; consideration sets
History: This paper was received January 25, 2005, and was with the authors 5 months for 2 revisions; processed by Eric Bradlow.

1. Noncompensatory Decision Processes

We explore new methods to study heuristic decision processes. These methods use "greedoid languages" and dynamic programming to solve combinatorial computational problems significantly more efficiently than as reported in the extant literature. We demonstrate how the methods can be used to identify the heuristic decision processes that best describe observed consideration and/or choice. Because the methods work with either rank-order data or consider-then-rank data, we are able to examine empirically how well each data-collection format predicts choice. A consider-then-rank task might be more enjoyable and less effortful for respondents (e.g., Malhotra 1986, Oppewal et al. 1994, Srinivasan and Park 1997) and, hence, might mean shorter questionnaires (less cost) and might encourage more respondents to complete the task (fewer nonresponse issues).

Noncompensatory decision processes are important both academically and managerially. Academically, there is ample evidence in the psychology, consumer behavior, and marketing science literatures that consumers simplify consideration and/or choice with a heuristic process.[1] Such processes are identified using a variety of methodologies ranging from verbal

[1] Examples include Bettman et al. (1998), Bettman and Park (1980), Bröder (2000), Einhorn (1970), Einhorn and Hogarth (1981), Gigerenzer and Goldstein (1996), Hauser (1978), Hauser and Wernerfelt (1990), Johnson and Meyer (1984), Luce et al. (1999), Martignon and Hoffrage (2002), Montgomery and Svenson (1976), Payne (1976), Payne et al. (1993), Roberts and Lattin (1991), Shugan (1980), Urban and Hauser (2004).

Figure 1 Cell Phones and SmartPhones

process tracing to information display mechanisms (e.g., Mouselab), and researchers have studied how consumers adapt and/or construct decision processes based on the characteristics of the decision environment (see Payne et al. 1993 for a review). Greedoid methods attempt to infer such processes from less intrusive observations where respondents are asked either to rank profiles, provide partial profile orders, or indicate whether or not they will consider a profile. These methods are, thus, complementary to existing methods.

Consider Figure 1 from a SmartPhone Web site[2] that encourages consumers to select SmartPhones for further consideration based on features such as carrier, brand, size, and price. Consumers can choose to keep or eliminate levels of these features to form a consideration set. If consumers use a heuristic, noncompensatory process and we can identify the features that drive the process, then the Web site designer knows which features to use, the product designer knows which features to include in the product line, and the advertising manager knows which features to emphasize.

Typically, compensatory conjoint analysis methods are used in these situations to identify the features with the largest partworths. Such methods are computationally tractable and often provide excellent paramorphic approximations of consumer consideration and/or choice processes. However, if consumers are not making compensatory trade-offs among features but, rather, are using noncompensatory heuristics, a method to identify those heuristics might be appealing. Heretofore, analyzing rank, partial rank, or consideration data to infer noncompensatory processes has not been computationally feasible for moderately sized problems because the number of potential noncompensatory descriptions grows as $n!$ where n is the number of distinct feature levels (aspects).[3] We provide an algorithm that can identify

[2] http://www.myphonefinder.com/, used with permission.

[3] Excellent methods exist for small numbers of aspects and for approximations to the best fit. Examples include Gensch (1987), Gilbride and Allenby (2004), Kohli and Jedidi (2004).

the best-fitting heuristic in seconds (rather than days). We illustrate the algorithm on two data sets and with experiments that vary the consumers' decision environment. Some of these experimental manipulations are designed to replicate existing findings; some experimental manipulations are new.

The paper proceeds as follows. First, we briefly review the literature on heuristic processes and provide examples. Next, we describe the respondents' tasks. We present greedoid-based methods and discuss the traditional methods to which they are compared. We test the methods empirically in a 2×2 experiment in which 339 respondents choose from 32 SmartPhones chosen from a fractional factorial $4^3 2^4$ design. We examine the impact of the number of profiles, the respondents' tasks, and sorting on the relative predictability of noncompensatory models. For comparison, we reanalyze classic data in which 201 respondents rated 16 computers chosen from a fractional factorial 2^{13} design. Finally, we close by illustrating how greedoid analysis provides managerial insight.

2. Brief Review of Noncompensatory Decision Processes

We consider decision processes in which products are represented by their features and consumers decide which product to purchase or consume. While the process by which consumers encode products into features can be complex and important (Einhorn and Hogarth 1981), that topic is beyond the scope of this paper. Our scope includes situations in which such encoding is feasible and reasonably descriptive of consumer decision processes. For practical applications, we might use voice-of-the-customer methods to identify a representative set of features (e.g., Griffin and Hauser 1993, Zaltman 1997). When a feature is binary, it is called an *aspect* (e.g., Tversky 1972). Multilevel features can be considered collections of aspects that are related (Verizon versus Cingular versus Nextel versus Sprint for SmartPhone service providers). A profile is the aspect description of a product.

Noncompensatory Processes

In a compensatory process, high levels on some aspects compensate for low levels on other aspects. In a noncompensatory process, high levels on some aspects cannot compensate for low levels on other aspects. One well-known noncompensatory process is a lexicographic process: Consumers evaluate profiles first by one feature, then another, until a judgment or choice is made (Fishburn 1974, Nakamura 2002). For example, consider an illustrative example in which consumers rank SmartPhones that differ on the *features* of brand and operating system. As illustrated by the first row of Figure 2, a consumer might rank first on the feature of brand, putting first all BlackBerry SmartPhones, then Nokias, Samsungs, and, last, Sony Ericsson, then rank on the feature of operating system (within brand) by putting all Microsoft-based SmartPhones before palm-based SmartPhones. We call this process lexicographic by features (LBF).

Other heuristics are possible. A consumer might rank SmartPhones by *aspects*, say, by first accepting BlackBerry SmartPhones, then Microsoft-based SmartPhones, Nokias, and, finally, Samsungs until all SmartPhones are ranked (second row of Figure 2). (Whenever there is a tie, the consumer moves to the next aspect in the lexicographic order.) For ease of reference, we call such processes *acceptance by aspects* (ABA). ABA is related to Tversky's elimination-by-aspects process (EBA) in which consumers successively eliminate aspects (third row of Figure 2). Tversky defines EBA as a random process in which the probability that an aspect is chosen is proportional to its measure. In this paper, we follow Johnson et al. (1989), Montgomery and Svenson (1976), Payne et al. (1988), and Thorngate (1980), and use EBA to refer to a deterministic process in which an aspect order is given. Finally, consumers may mix acceptance and elimination criteria. We call such a mixed process *lexicographic by aspects* (LBA). Because ABA, EBA, and LBF are special cases of LBA, we focus on LBA.

When a feature has more than two aspects, eliminating an aspect (Sony Ericsson) is the same as accepting its complement (BlackBerry $\cup$ Nokia $\cup$ Samsung), but EBA is not equivalent to an ABA process of accepting BlackBerry, then Nokia, then Samsung. The ABA process orders BlackBerry-Nokia-Samsung, while the EBA process does not. However, for two-level aspects, there exists an equivalent EBA process for every ABA process. Figure 3 illustrates this equivalency for SmartPhones that differ on three binary aspects (brand, operating system, and carrier).

ABA, EBA, LBA, and LBF define orderings and hence can be used to explain either full or partial rankings, including respondent tasks such as rank all profiles, choose a single profile, or indicate which profiles are worth further consideration. Finally, the processes can be modified to include constraints within features such as "lower prices are always preferred to higher prices."

Compensatory Processes

Many authors represent a compensatory process as an arithmetic rule in which each aspect receives a weight and consumers sum the weights associated with the aspects in a profile to form "utility." Consumers then choose the product with the highest "utility." However, not all sets of aspect partworths imply a compensatory process. If the aspect partworths follow

Figure 2 **Examples of Lexicographic Heuristic Processes**

Simplifying heuristic	TLA (*Three-letter acronym*)	Ranking rule
Lexicographic by features	LBF	(BlackBerry > Nokia > Samsung > SONY Ericsson), (Microsoft > PalmOS)
Acceptance by aspects	ABA	BlackBerry, Microsoft, Nokia, Samsung
Elimination by aspects	EBA	SONY Ericsson, PalmOS, Samsung, Nokia
Lexicographic by aspects	LBA	SONY Ericsson, BlackBerry, Microsoft, Nokia

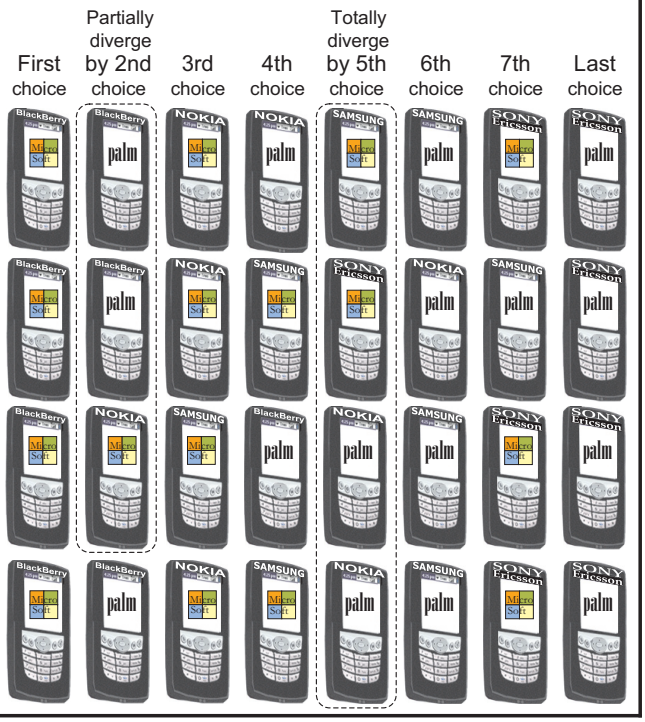
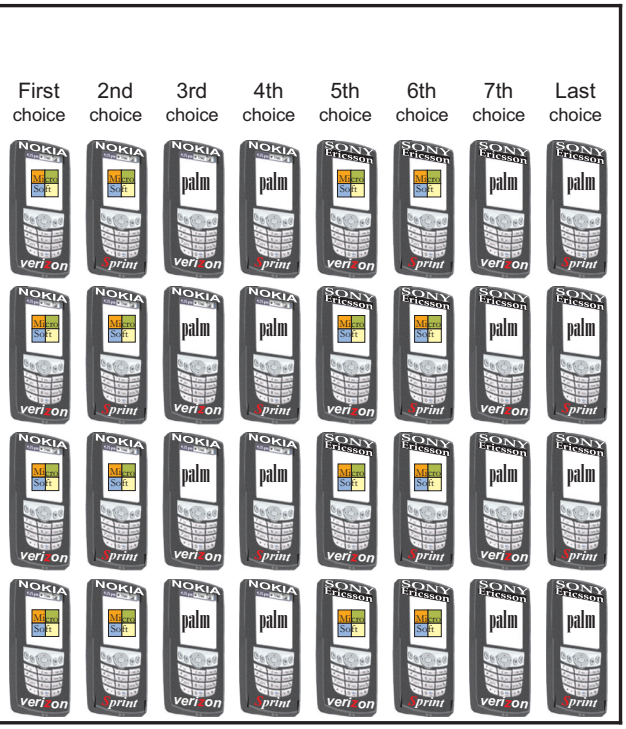


Figure 3 **Equivalent Processes for Binary Features (Two-Aspect Features)**

Simplifying heuristic	TLA (*Three-letter acronym*)	Ranking rule
Lexicographic by features	LBF	(NOKIA > SONY Ericsson), (Microsoft > PalmOS), (verizon > Sprint)
Acceptance by aspects	ABA	NOKIA, Microsoft, verizon
Elimination by aspects	EBA	SONY Ericsson, PalmOS, Sprint
Lexicographic by aspects	LBA	SONY Ericsson, Microsoft, verizon



an appropriate geometric sequence (e.g., 2^{1-n} for the nth aspect), then an additive model produces a lexicographic process in which no set of lower ranked aspects can compensate for the lack of a higher ranked aspect (Jedidi et al. 1996, Kohli and Jedidi 2004, Olshavsky and Acito 1980). Thus, we reserve the word "compensatory" for additive models that are truly compensatory, e.g., when the partworths are constrained so that the presence of other aspects can compensate for the lack of an important aspect.

Constructive Processes

Research suggests that consumer decision processes are contingent on many context effects including the range of aspects, correlation among aspects, base-rate information, reference points, the size of the choice set, the relevance of the decision, and the difficulty of comparison (see review in Payne et al. 1993). To the extent that data collection approximates the essential characteristics of real choice environments, greedoid methods provide insight into how context affects respondents' tendency to use noncompensatory decision rules. Our empirical experiments illustrate this context-dependent variation.

Existing Methods to Infer Noncompensatory Processes

Many measurement examples in the marketing science literature are consistent with noncompensatory decision processes. For example, both Srinivasan and Wyner's (1988) Casemap and Johnson's (1991) adaptive conjoint analysis (ACA) include steps in which respondents are asked to eliminate unacceptable levels. Because this task is often difficult for respondents (Green et al. 1988, Klein 1988), other researchers have attempted to infer the elimination process in a single estimation step (DeSarbo et al. 1996, Gilbride and Allenby 2004, Gensch 1987, Gensch and Soofi 1995, Jedidi and Kohli 2005, Jedidi et al. 1996, Kim 2004, Roberts and Lattin 1991, Swait 2001). For example, Gilbride and Allenby (2004, p. 399) use hierarchical Bayes methods to analyze choice-based conjoint data to infer screening rules for cameras. They estimate that 58% of the respondents screen on a single feature, 33% on two features, 2% on three features, and 8% use fully compensatory processes.

In psychology, Bröder (2000) analyzes choices among two profiles described by four aspects. He compares the fit of an unconstrained additive model to two additive models: (1) a model in which the aspects are constrained to 2^{1-n} (noncompensatory), and (2) a model with equal weights (Dawes' 1979 model). In one experiment, 28% of the 40 respondents are classified as noncompensatory while none are classified as Dawes'. The remaining 72% could not be classified. Bröder's (2000) method is feasible for a small number of aspects—with four aspects, the ratio of the largest-to-smallest partworth in a noncompensatory model is $2^3 \rightarrow 8{:}1$. For 16 aspects, as in our experiments, the range of partworths in a noncompensatory model would be at least $2^{15} \rightarrow 32{,}768{:}1$, a ratio that puts severe strains on any statistical regression-like procedure.[4]

Kohli and Jedidi (2004) propose a greedy heuristic to estimate a linear representation of lexicographic processes from metric conjoint data.[5] They modify Bröder's (2000) procedure by computing the number of violated pairs between predicted and observed rank orders for both additive and 2^{1-n} models. Their t-statistics suggest that a 2^{1-n} representation is not significantly different from an unconstrained additive model for 67% of the 69 respondents who evaluated profiles with 5 features (11 aspects). Our approach differs from Kohli and Jedidi (2004) along a number of dimensions including respondent task, estimation algorithms, and focus. Nonetheless, these parallel independent studies suggest many opportunities to apply discrete optimization methods to infer noncompensatory processes.

3. The Respondents' Tasks

It is easier to understand greedoid methods if we first review the respondents' tasks as illustrated with an example from our empirical experiments.[6] In Figure 4, respondents are first introduced to the product category and the 7 features (16 aspects). Figure 4a is one of many screens. Respondents are then presented with SmartPhone profiles (Figure 4b). Respondents in the consider-then-rank cells simply click on those profiles they would seriously consider—part of the screen is shown in Figure 4c. These respondents then see only their considered profiles; they are asked to rank them by successively clicking on the profile they would choose from the offered set (Figure 4d). That profile disappears and they choose again until all considered profiles are chosen. Respondents in the full-rank cells skip the consideration task. In Figure 4, we illustrate an additional twist. Some, but not all, respondents were allowed to presort the profiles in either or both tasks. A priori, we expect the ability to sort tasks to encourage lexicographic processing. In a later section, we describe the full experimental protocol (incentives,

[4] The ratio is not as severe if we allow partial lexicographic orders. However, Bröder's (2000) method cannot handle partial orders without first solving the combinatorial problems we describe later.

[5] Both approaches were developed independently. We became aware of one another's approaches after all empirical work had been completed and papers written.

[6] Greedoid analysis can be extended to tasks such as those used by Bröder (2000), Gigerenzer and Goldstein (1996), or Gilbride and Allenby (2004).

Figure 4 Illustrative Respondent's Task

(a) Describing aspects to respondents

(b) Example profiles

(c) Consideration judgment task

(d) Choice from consideration set

Note. These figures have been modified from the originals to avoid the prominent use of corporate logos.

filler tasks, holdout measures, etc.) and provide statistics such as response rates.

Two comments are in order. First, respondents are *not* asked to indicate unacceptable feature levels (aspects) directly. Elimination aspects, if any, are inferred from the data. Second, the measurement tasks themselves do not assume that either judgment (consideration) or decision (choice or rank) processes are compensatory or that they are noncompensatory.

4. Identifying Lexicographic Processes with Greedoid Languages

For ease of exposition, we focus our discussion on lexicographic aspect orders that represent acceptance-by-aspects (ABA) decision rules. Elimination by aspects (EBA) can be estimated with the same algorithm by redefining aspects by their negation; lexicographic by features (LBF) can be estimated by imposing constraints on the aspect orders; and lexicographic by aspects (LBA) can be estimated with only a slight modification in the algorithms.

To identify the best lexicographic representation for a given set of observed data, we develop a procedure to identify the aspect order that maximizes fit (minimizes errors) on some metric. Unfortunately, as Martignon and Hoffrage (2002, Theorem 2, p. 39) demonstrate, this problem is NP-hard. They suggest exhaustive enumeration. For example, they sought to determine the aspect order (state capital, soccer team in the national league, etc.) that best explains the relative populations of pairs of German cities. Because their problem had nine aspects, they needed to search 9! orderings—"a UNIX machine" took two days to find the best ordering. Their problem is a relatively small problem. For our $4^3 2^4$ empirical problem, exhaustive enumeration needs to check $7! \times (4!)^3$ orders for LBF, 16! aspect orders for ABA or EBA, and $2^{16} \times 16!$ orders for LBA.[7] Because $7! \times (4!)^3 = 192 \times 9!$, their algorithm would have taken over a year per respondent for LBF. Because $16! = 300{,}300 \times 7! \times (4!)^3$, their algorithm would have taken over 300 millennia for ABA or EBA and substantially longer for LBA.

[7] In a $4^3 2^4$ design, there are $3 \times 4 + 4 \times 1 = 16$ aspects.

Although faster computers help, it is clear that practical analysis for moderate-to-large problems requires a more efficient algorithm.

We address computational efficiency in two steps. We first demonstrate that the collection of lexicographic orders of *aspects*, as they relate to a partial ordering of *profiles*, forms a "greedoid language."[8] This enables us to use established results to find the appropriate lexicographic aspect order, if one exists, much more efficiently than existing methods. Because the profile ordering need only be partial, we can handle consideration, partial-rank, full-rank, or choice data.

A perfect lexicographic ordering of aspects is extremely rare. With 32 profiles, there are 32! rank orders but only $2^{16} \times 16!$ aspect orders. Thus, the chances that an arbitrary profile order is consistent with an aspect order is less than 5.2×10^{-18}. Furthermore, any respondent errors are likely to cause the data to be inconsistent with an aspect order.[9] Using the greedoid structure, we prove that a dynamic program can find the lexicographic ordering that maximizes a commonly used goodness-of-fit metric. The dynamic programming algorithm substantially reduces computation and makes it feasible to identify the best lexicographic ordering for large samples of respondents and moderately large numbers of aspects. For example, for 16 aspects, the dynamic program need only consider $2^{16} = 65{,}536$ subsets of aspects, much less than the 1.4×10^{18} potential LBA aspect orders. We begin with notation and definitions.

Partial Orders and Consistency

Let N be the total number of aspects, and let $L = (L_1, \ldots, L_a)$ be an ordered subset of a aspects where $a \leq N$. For a given profile P, we let $L_i(P)$ be one if profile P contains aspect L_i, and zero otherwise. We write $P \succ_L P'$ if $L_i(P) = 1$ and $L_i(P') = 0$, where i is the first (smallest) index for which $L_i(P) \neq L_i(P')$. For example, in the ABA row of Figure 2, $L =$ (BlackBerry, Microsoft, Nokia, Samsung). The first ranked SmartPhone is $P = \{\text{BlackBerry, Microsoft}\}$. $L_1(P) = 1$, $L_2(P) = 1$, $L_3(P) = 0$, and $L_4(P) = 0$. $P = \{\text{BlackBerry, Microsoft}\} \succ_L P' = \{\text{BlackBerry, PalmOS}\}$ because $L_2(P) = 1$ and $L_2(P') = 0$.

Each totally ordered set L of aspects implies a unique order of profile preferences, but the converse is not true. Different orders of aspects can lead to the same order of profiles. This is particularly true for the consideration task. If a respondent will consider only Verizon SmartPhones that flip open, then the orders $L =$ (Verizon, flip) and $L' =$ (flip, Verizon) are both consistent with the respondent's consideration process.

Suppose that X is a partial order of profiles revealed by the respondent. For example, X might define which profiles are in a consideration set and which are not, or X might define a rank order within a consideration set. We write $P \succ_X P'$ if profile P is preferred to (or ranked higher than) profile P' according to X. We say that an ordered subset L of aspects is *lexico-inconsistent* with a partial order X of profiles if there are profiles P and P' that are ranked differently by aspect order L and profile order X. In symbols, L and X are inconsistent if there exist P and P' such that $P \succ_X P'$ while $P' \succ_L P$. Otherwise, we say that L and X are *lexicoconsistent*. For example, in the ABA row of Figure 2, the aspect order in the ranking rule column (BlackBerry, Microsoft, Nokia, Samsung) is lexicoconsistent with the order of the eight profiles in the final column.

If L is an ordered subset of aspects and $e \notin L$ is an aspect, then let (L, e) denote the ordered subset of aspects obtained by appending aspect e to the end of L. Let $L \backslash Y$ denote the set L with all elements of Y deleted. For example, if $L =$ (BlackBerry, Microsoft, Nokia, Samsung), $e =$ flip, and $Y =$ (Samsung), then $(L, e) =$ (BlackBerry, Microsoft, Nokia, Samsung, flip) and $L \backslash Y =$ (BlackBerry, Microsoft, Nokia). Finally, let $\varnothing$ be the empty set, and let the number of elements in a set L be denoted by $|L|$.

Greedoid Languages

Greedoid languages were developed by Korte and Lovász (1985) to study conditions under which a greedy algorithm can solve optimization problems.[10] They have proved useful in sequencing and allocation problems (e.g., Niño-Mora 2000). We believe that this is the first application in marketing or consumer behavior. Björner and Ziegler (1992) and Korte et al. (1991) are excellent references that provide numerous examples of greedoids. In this work, we introduce a new type of greedoid where partial orders of profiles (X) induce a greedoid language defined on aspect orders (L). It is this linkage that enables us to identify LBA and other aspect- or feature-based explanations of profile orders.

Let E be a set of aspects and let G be a collection of ordered subsets of E. We say that G is a greedoid language if the following conditions are satisfied:

[8] Early theory examined partially ordered set greedoids such as might be defined on orderings of profiles. The greedoid in this paper describes aspects as they relate to profiles, not profiles directly.

[9] Interestingly, less that one-tenth of 1% of the profile orderings are consistent with linear combination of aspect measures. This includes both compensatory and noncompensatory linear combinations.

[10] Greedoids were initially formulated in the context of set systems. However, the language representation is equivalent and more appropriate for our application.

1. $\varnothing \in G$.

2. If $L \in G$, and if element $e \in E$ is the last element of L, then $L\backslash e \in G$.

3. If $L \in G$ and if $L' \in G$ and if $|L'| > |L|$, then there is an element e in $L'\backslash L$ such that $\hat{L} = (L, e)$ and $\hat{L} \in G$.

In the appendix, we demonstrate that the set of partial orderings of the aspects that are consistent with a partial ordering of the profiles form a greedoid language. Corollary 1 follows from Proposition 1 because Algorithm 1 is a greedy algorithm that either finds a lexicoconsistent aspect order, $L \in G$, of maximum length or terminates early.[11]

Proposition 1. *Let E be the set of aspects, and let X be a partial order on the profiles. Let G be the collection of ordered subsets of E that are lexicoconsistent with X. Then, G is a greedoid language.*

Corollary 1. *Algorithm 1 determines whether there exists a lexicographic ordering of aspects L that is lexicoconsistent with a profile ordering X and, if an ordering exists, finds an ordering.*

Algorithm 1 (for determining if L is lexicoconsistent with X).

```
begin
  L = ∅
  while |L| < |E| do
  begin
    if (there exists e in E\L such that (L, e) is
        lexicoconsistent with X)
      replace L with (L, e)
    else
      quit (because X is not lexicographic)
  end
end
```

Finding Lexicographic Descriptions that Maximize Fit

If there is no ordering of aspects that is lexicoconsistent with a profile order, i.e., Algorithm 1 terminates with $|L| < |E|$, we might still like to find ordering(s) of aspects that best fit a profile order. As a measure of fit between aspect orders (L) and profiles (X), we relate "closeness" to the number of inconsistencies (violated pairs) between the profile order induced by L and that observed in X.[12] We now develop an algorithm to find the closest lexicographic ordering. We begin with Proposition 2, which explores the implications of the greedoid structure. Proposition 2 implies Algorithm 2, which is a dynamic program on the set of all subsets of the aspects. Because this set has dimensionality 2^n, and because 2^n is substantially less than $n!$, Algorithm 2 can still be feasible when exhaustive enumeration is not.

In the appendix, we prove Proposition 2 by showing that the marginal inconsistencies induced by adding a new aspect to an existing aspect order depend only on that aspect and the set of aspects preceding it, but not on the *order* of the preceding aspects. This implies a (forward) recursive structure that allows the problem to be solved with dynamic programming (Corollary 2). This dynamic program is similar to Held and Karp's (1962) classic dynamic programming algorithm for the traveling-salesman problem. For those readers not familiar with dynamic programming, we provide a technical appendix (available at http://mktsci.pubs.informs.org on the *Marketing Science* Web site) that illustrates how Algorithm 2 would apply to a simple playing card example.

Proposition 2. *Let L and L' be two different permutations of a subset E' of aspects, and let e be any aspect not in E'. Then, the number of lexico-inconsistencies directly caused by e in (L, e) is the same as the number of inconsistencies caused by e in (L', e).*

Corollary 2. *Algorithm 2 identifies the lexicographic ordering of aspects L that best fits the (partial) ordering of profiles X.*

When Algorithm 2 terminates, $J(E)$ is the minimum number of lexico-inconsistencies between the respondent's profile ordering X and any ordering of the aspects in E. $L(E)$ are the best-fitting lexicographic orders, which might or might not be unique. Algorithm 2 applies directly to either acceptance by aspects or elimination by aspects. Fortunately, for lexicographic by aspects, the number of steps in the algorithm only doubles. This is a significant improvement relative to exhaustive enumeration, which would cause the number of steps to grow by a factor of 2^n for LBA. Specifically, in the innermost loop of Algorithm 2 we need only check both i and its negation. We call Algorithm 2 a greedoid-based dynamic program.

Algorithm 2 (for finding aspect order L that provides the best fit to profile order X).

```
begin
  J(∅) = 0
  for k = 1 to |E|
    for all (unordered) subsets, S ⊆ E of size k
      for all i ∈ S
        c(S\{i}, i) = number of inconsistencies caused
          by aspect i following set S\i
      next i
      J(S) = min_{i∈S}[J(S\{i}) + c(S\{i}, i)]
      L(S) is the ordering of aspects in S yielding
        J(S) [retained]
```

[11] Kohli and Jedidi (2004) use a greedy algorithm on permutation matrices to identify a metric representation of a lexicographic ordering when there is no response error. They apply their algorithm to metric data.

[12] Minimizing violated pairs is equivalent to maximizing Kendall's tau (1975) where $\tau = 1 - 2 * (\text{fraction violated})$.

```
    next S
  next k
end
```

We programmed Algorithm 2 in Java running on an IBM 1.7-GHz laptop. For a 16-aspect problem, the run time was approximately 1.85 seconds per respondent. Relative to Martignon and Hoffrage (2002), some savings are due to Algorithm 2 and some are due to faster computers. We project that Martignon and Hoffrage's (2002) exhaustive enumeration would take 14 years for a 16-aspect EBA problem on the same computer.[13]

We provide two further results. Proposition 3 extends the theory (and Algorithms 1 and 2) to allow the researcher to place greater emphasis on some ordered pairs in X, say, the ordered pairs corresponding to highly ranked profiles. Proposition 4 reduces the running time of Algorithm 2 by enabling it to begin with any consistent ordered subset of aspects (e.g., the largest such ordered subset) and then use the dynamic program on the remaining aspects.

Proposition 3. *If weights are associated with each ordered pair in X, then* (1) *the new G is a greedoid language*, (2) *Algorithm* 1 *determines whether there exists an L lexicoconsistent with X*, (3) *Proposition* 2 *extends to the new G, and* (4) *Algorithm* 2 *finds the best fit if $c(\cdot)$ is redefined to $c(S\backslash\{i\}, i) =$ sum of weighted violations caused by aspect i following set $S\backslash i$.*

Proposition 4. *Suppose that L is an ordered subset of aspects that is lexicoconsistent with the preferences of X. Then, there is an optimal ordering of aspects that begins with the order L.*

5. Benchmarks

To evaluate the ability of the greedoid methods to fit partial- or full-rank data and predict holdout validations, we identify benchmark models. Because hierarchical Bayes methods appear to be the most popular method to estimate additive models, our primary comparison is a hierarchical Bayes-ranked logit model (HBRL, e.g., Rossi and Allenby 2003). We provide an alternative benchmark with linear programming estimation (LINMAP). We use the most recent version of LINMAP, which enforces strict rankings (Srinivasan 1998). Both benchmark methods predict holdouts slightly better than either traditional LINMAP (Srinivasan and Shocker 1973) or analytic center estimation (Toubia et al. 2003).[14]

With HBRL (or LINMAP) we must address the conceptual issue that additive models nest lexicographic models. The best-fitting additive model might estimate partworths that are equivalent to a lexicographic process. For example, estimated partworths that satisfy a 2^{1-n} relationship have this property. This would be fine if our only interest were predictive ability. However, we also seek to use greedoid methods to gain insight on consumers' heuristic processes.

To estimate additive models that do not nest lexicographic models, we constrain the additive model so that all estimated partworths are truly compensatory. By the principle of optimality, such a constraint cannot be estimated by fitting data.[15] On the other hand, such a constraint might improve *holdout* performance by the principle of complexity control (Cui and Curry 2005, Evgeniou et al. 2005).

The form of our constraints is motivated by behavioral researchers who have sought to identify whether compensatory or noncompensatory models fit or predict observed choices better. For example, Bröder (2000) defines a respondent as compensatory if the respondent's partworths are "not too extreme." Specifically, Bröder requires that $w_{ic} = w_{lc}$ for all $l \neq i$, where w_{ic} is the partworth of the ith aspect for respondent c. We generalize Bröder's (2000) precedent by defining a respondent as "q-compensatory" if $w_{ic} \leq q w_{lc}$ for all $l \neq i$. With this definition, we can examine a continuum between Dawes' (1979) model as tested by Bröder ($q = 1$) and the unrestricted additive benchmark ($q = \infty$) that nests lexicographic models. Because there is no a priori theory with which to select q, we provide holdout predictions for values of q ranging from 1 to ∞. We estimate q-compensatory benchmarks with rejection sampling in the hierarchical Bayes sampler or by imposing additional constraints on the linear program. We label these q-compensatory benchmarks HBRL(q) and LINMAP(q). In a supplemental appendix (available on the *Marketing Science* Web site), we use synthetic data to explore the implications of q for "true" models that vary from highly compensatory to highly lexicographic. For these simulations, selecting $q = 4$ provides a reasonable ability to discriminate respondents with "compensatory" partworths from those with "lexicographic" partworths.

6. SmartPhone Empirical Study

To test greedoid methods, we invited respondents to complete a web-based questionnaire about SmartPhones. The respondents were students drawn from the undergraduate and graduate programs at two universities. To the best of our knowledge, they were unaware of greedoid methods or the purpose of

[13] Exhaustively enumerating LBA would take over 900 millennia. Exhaustively enumerating a nine-aspect problem would take nine seconds for EBA (or ABA), but 1.3 hours for LBA.

[14] Details on classical LINMAP and analytic-center estimation are available from the authors.

[15] Unconstrained models necessarily fit better than the constrained models that they nest; thus, fit cannot be used as a criterion.

our study. As an incentive to participate, they were offered a one-in-ten chance of winning a laptop bag worth $100, yielding a 63% response rate. Pretests in related contexts suggested that SmartPhones were likely to include noncompensatory features and thus represented an interesting category for a first test of greedoid methods.

The survey consisted of six phases. The first three phases are as described in Figure 4: Respondents reviewed the category and SmartPhone features, indicated which SmartPhones they would consider (in half the cells), and successively chose SmartPhones in order to rank their considered products (or rank all products, depending on cell). Respondents then completed a mini-IQ test to cleanse memory—a task which pretests suggested was engaging and challenging. Following this filler task, respondents completed a holdout task consisting of two sets of four SmartPhones chosen randomly from a different 32-profile fractional factorial design.[16] The final task was a short set of questions about the survey itself—data which we use to compare task difficulty. For the holdout task, to avoid unwanted correlation due to common measurement methods, we used a different interface. Respondents used their pointing device to shuffle the profiles into a rank order as one might sort slides in PowerPoint. Pretests suggested that respondents understood this task and found it different from the task in Figure 4d.

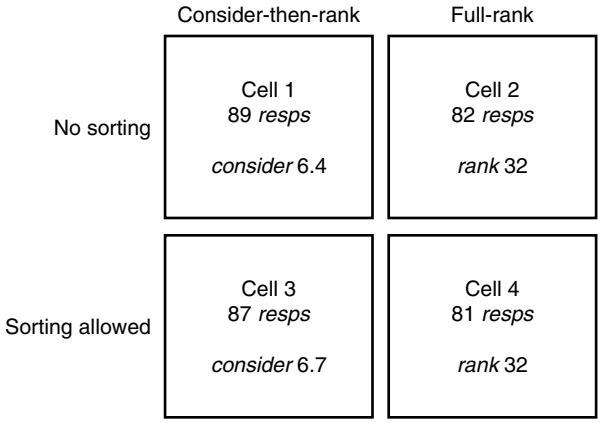
The survey was programmed in PHP and debugged through a series of pretests with 56 respondents chosen from the target population. By the end of the pretests, all technical glitches were removed. Respondents understood the tasks and found them realistic.

Experimental Design

Respondents were assigned randomly to experimental cells. The basic experimental design is a 2×2 design in which respondents complete either a full-rank or a consider-then-rank task and are given the opportunity to presort profiles or not (Figure 5). In the consider-then-rank sort cell, respondents could sort prior to consideration *and* prior to choice. Respondents in the sort cells could re-sort as often as they liked. We also included an additional cell (described below) to test whether the results vary by the number of profiles presented to the respondents in the consider-then-rank cells. This experimental design enables us to test greedoid methods with different data collection tasks and to illustrate how greedoid methods might be used to explore how context affects respondents' processing strategies.

[16] Future research might investigate the effect of wear-out on lexicographic processing with cells that place the holdout tasks earlier in the survey. See also Hauser and Toubia (2006) and Liechty et al. (2005).

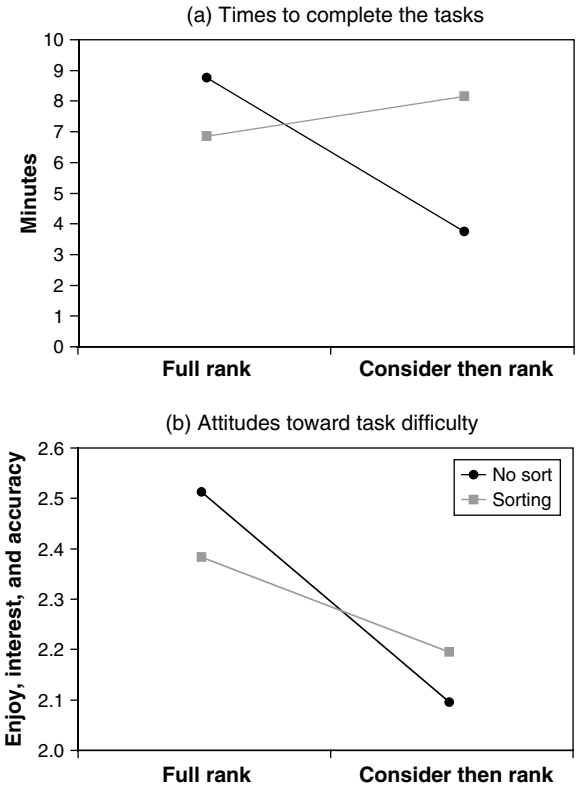
Figure 5 SmartPhone Experimental Design (32 Profiles in $4^3 2^4$ Fractional Design



Task Difficulty

Greedoid methods can be used to analyze any full- or partial-order respondent task. We first examine whether the consider-then-rank task is more natural and easier for respondents than the full-rank task. The results are reported in Figures 6a and 6b. We oriented both axes such that down is better. In the base condition of no sorting, the consider-then-rank task is seen as significantly more enjoyable, accurate, and engaging ($t = 2.2$, $p = 0.03$), saves substantial time (3.75 minutes compared to 8.75 minutes,

Figure 6 Task Difficulty (Less Is Better on Both Graphs)



$t = 2.8$, $p = 0.01$), and appears to increase completion rates (94% versus 86%, $t = 1.7$, $p = 0.09$). Sorting (as implemented) mitigates these advantages: Neither attitudes, time, nor completion rates are significantly different between the full-rank and consider-then-rank tasks when respondents can presort profiles.[17] A possible explanation is that sorting made the full-rank task "easier" (though not necessarily more enjoyable) and the consider-than-rank task more complex.

Predictive Ability

We first compare the most general greedoid method (LBA) to the unconstrained additive models HBRL and LINMAP, as averaged across respondents (see Table 1). Holdout predictions are based on two metrics. Hit rate provides fewer observations per respondent (two) and leads to more ties, but it is not optimized directly by either greedoid methods or the benchmarks. The percent of violated pairs provides more observations per respondent (12 potential pairs from two sets of four ranked profiles), but it is the metric optimized by greedoid methods and, to some extent, by LINMAP. Empirically, the two metrics are significantly correlated (< 0.001 level) for all methods and provide similar comparative qualitative interpretations.[18]

As expected, the unconstrained LINMAP, which nests LBA and optimizes a metric similar to the fit metric, provides the best fit. However, LBA fits almost as well. The more interesting comparisons are on the two holdout metrics. For both metrics, LBA is better than both benchmarks and significantly better on hit rates. It appears that, for these data, greedoid methods are more robust than the unconstrained additive models that could, in theory, fit a lexicographic process. This apparent robustness is consistent with predictions by Mitchell (1997) and Martignon and Hoffrage (2002, p. 31). We address the last column of Table 1 later in this section.

Comparison to q-compensatory Processes

Following Bröder (2000), we examine whether respondents are described better by lexicographic or q-compensatory processes. Three comments are in order. First, this description is paramorphic. We say only that respondents rank (choose, consider) profiles *as if* they were following one or the other process. Second, we have some confidence in the descriptions because LBA predicts better for synthetic respondents who are lexicographic, and a constrained additive model (q-compensatory) predicts better for synthetic respondents who are q-compensatory (see supplemental appendix available on the *Marketing Science* Web site). Third, for simplicity of exposition, we compare LBA to the HBRL benchmark. This benchmark does slightly better than LINMAP in Table 1 and, as we will see later, better for a second data set. Comparisons to LINMAP are in a supplemental appendix (available on the *Marketing Science* Web site).[19]

Table 1 Comparison of Fit and Prediction for Unconstrained Models

	Lexicographic by aspects	Hierarchical Bayes ranked logit	LINMAP	Lexicographic by features
Fit (percent pairs)	0.955*	0.871	0.969†	0.826
Holdout percent pairs	0.745*	0.743	0.737	0.658
Holdout hit rate	0.597**	0.549	0.549	0.481

*LBA significantly better than LBF. **LBA significantly better than HBRL, LINMAP, and LBF. †LINMAP significantly better than LBA and HBRL. Tests at the 0.05 level.

Figure 7 plots holdout predictions as a function of q. Predictions improve as the models become less constrained (larger q), consistent with a perspective that some aspects are either being processed lexicographically or have large relative partworths. HBRL(q) approaches LBA's holdout percent pairs predicted for large q but falls short on holdout hit rates.

At the level of the individual respondent, comparisons depend upon the choice of q. As an illustration, we use $q = 4$. When $q = 4$, the respondent is acting as if he or she is making trade-offs among aspects by weighing their partworths. Furthermore, the analysis of synthetic data suggests that at $q = 4$ respondents who are truly compensatory are classified as compensatory and respondents who are truly lexicographic are classified as lexicographic.

For holdout percent pairs, LBA predicts better than HBRL(4) for 56% of the respondents, worse for 43% of the respondents, and tied for 1% of the respondents. On average, LBA's predictive ability is about five percentage points higher than HBRL(4). The corresponding comparative percentages for hit rates are 46%, 30%, and 24%.[20] On average, LBA's hit rate is about 11 percentage points higher than HBRL(4). Figure 8 provides a visual comparison of the distributions of holdout metrics for individual respondents. Positive numbers (darker bars) indicate those respondents for which LBA predicts better than HBRL(4). These percentages and Figure 8 suggest that greedoid methods

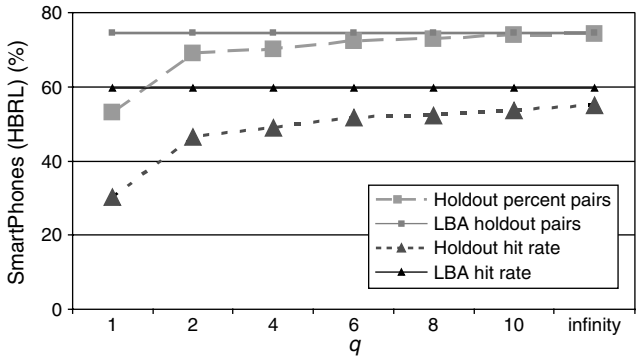
[17] For the sorting cells, attitudes ($t = 0.9$, $p = 0.37$), time ($t = 0.4$, $p = 0.70$), and completion rate ($t = 1.1$, $p = 0.26$) are not significantly different. Using analysis of variance, there is an interaction between sorting and task for time, but it is not significant ($F = 2.6$, $p = 0.11$). For attitudes only, task is significant ($F = 4.9$, $p = 0.03$).

[18] For example, correlations between the metrics are 0.70 for LBA, 0.64 for HBRL, and 0.66 for LINMAP.

[19] For the SmartPhone data, for some values of q, LINMAP does better than HBRL. The relative performances of the benchmarks are interesting but beyond the scope of this paper.

[20] At the level of individual respondents, hit rates are coarser measures than the percent of violated pairs, hence more ties are observed.

Figure 7 Comparison of Holdout Prediction for q-compensatory Models



are a viable method to complement more traditional methods to evaluate whether respondents are using compensatory or noncompensatory processes.

Constructed Processes—Full Rank vs. Consider Then Rank; Sorting vs. Not Sorting

Behavioral researchers hypothesize that consumers construct their decision processes as they make their decisions and hence that these decision processes can be influenced by the nature of the decision task. We examine this issue by comparing the influence of task (consider then rank versus full rank) and the availability of a presorting mechanism (sorting allowed versus not allowed). Figure 9 compares the predictive ability (holdout violations) for the four cells of our basic experiment. Some insights from Figure 9 are:

- Allowing respondents to presort SmartPhones does not have a significant effect on either LBA or HBRL(4). Task has a significant effect on both LBA and HBRL(4).[21]
- On average, LBA predicts significantly better than a q-compensatory model in full-rank cells ($t = 6.0$, $p = 0.0$) but not in the consider-then-rank cells ($t = 0.4$, $p = 0.69$).
- A lexicographic model predicts better than a q-compensatory model for more respondents in the

Figure 8 Histograms of Comparative Predictive Ability

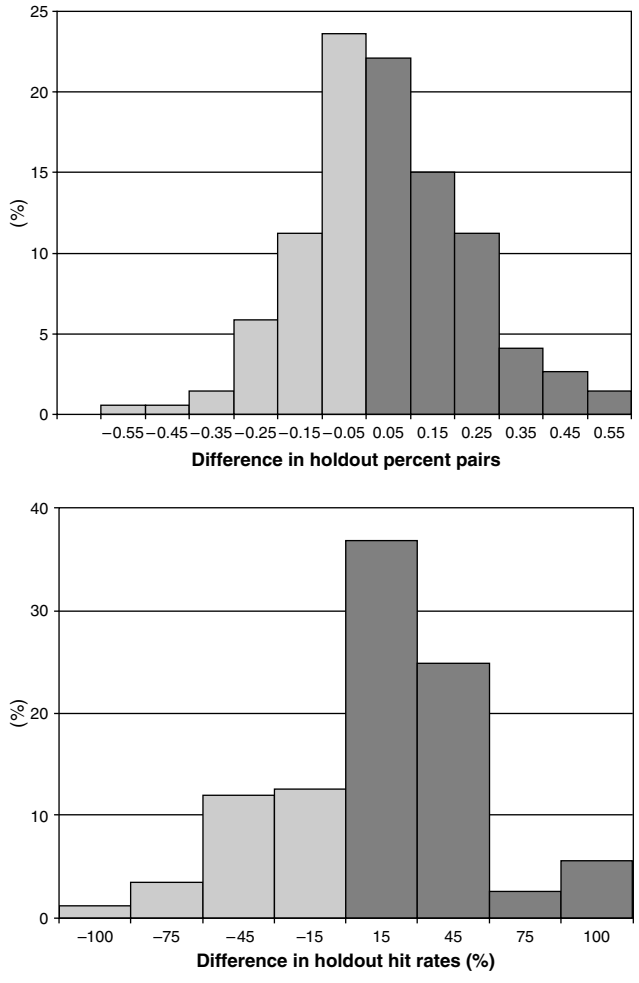
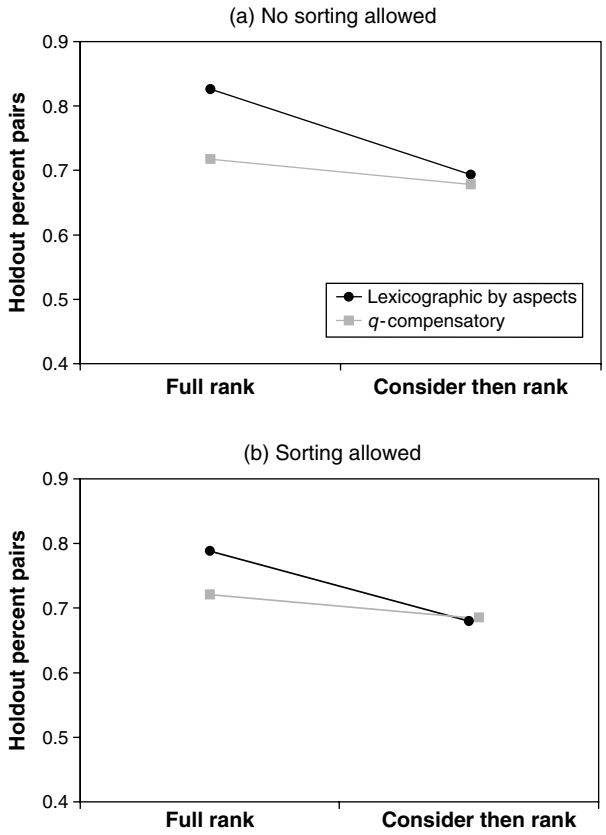


Figure 9 Predictive Ability by Experimental Cell, Lexicographic vs. q-compensatory Processes



[21] Using analysis of variance, task is significant for both LBA ($F = 51.1$, $p = 0.00$) and HBRL(4) ($F = 3.7$, $p = 0.05$). Sorting is not significant for either LBA ($F = 2.1$, $p = 0.14$) or HBRL(4) ($F = 0.1$, $p = 0.79$).

full-rank cells than in the consider-then-rank cells (62% versus 50%, $t = 2.2$, $q = 0.03$).[22]

We obtain a similar pattern of results for hit rates, with the exception that hit rates are a coarser measure at the level of the individual respondent (more ties) and require a relative measure.[23]

Constructed Processes—Predictive Ability vs. Effort

Data in the previous section are consistent with a hypothesis that the more effortful experimental cells (full rank versus consider then rank) lead to more lexicographic processing. We can also manipulate effort by the number of profiles that the respondent is asked to evaluate. Indeed, behavioral theory suggests that respondents are more likely to use a lexicographic process *for choice* (rank) if there are more profiles (e.g., Bettman et al. 1998, Johnson et al. 1989, Lohse and Johnson 1996). Payne et al. (1993, pp. 253–254) suggest as important research the study of such manipulations on consideration set formation.

To examine this issue, we assigned an additional 86 respondents to a fifth cell in which respondents evaluated fewer profiles (16 versus 32) using the *consider-then-rank* task. With this manipulation, we found no significant differences in the relative predictive ability of LBA versus HBRL(4) between cells ($t = 0.2$, $p = 0.88$ for percent pairs predicted, and $t = 1.0$, $p = 0.31$ for the percent of respondents for whom LBA predicts better). We obtain the same pattern of results with hit rates. Interestingly, the differences in effort are also not significant for 16 versus 32 profiles when the task is consider then rank.[24] Perhaps the number of profiles has less of an effect on consideration than that reported in the literature for choice—an empirical result worth examining in future experiments. Alternatively, the 16-profile task might have already been sufficiently difficult to trigger the use of simplifying heuristics for consideration.

We did not include a cell in which respondents were asked to provide full ranks for 16 profiles. However, to gain insight, we simulate a 16-profile full-rank cell by randomly choosing one-half of the 32 profiles for estimation. Predictions degrade with half the profiles, but the loss is less than three percentage points (80.8% versus 77.9%, $t = 4.3$, $p = 0.00$).[25]

The effect of task type seems to have a larger impact than the number of profiles. LBA estimates from the full-rank task predict significantly better than those from the consider-then-rank task (review Figure 9). On average (combining sort and no-sort cells), 81% of the holdout pairs are predicted correctly in the full-rank cells compared to 69% in the consider-then-rank cells ($t = 2.6$, $p = 0.01$). On the other hand, the consider-then-rank task took significantly less time to complete in the no-sort cell ($8\frac{3}{4}$ versus $3\frac{3}{4}$ minutes).

The three effort comparisons (full rank versus consider then rank, 16 versus 32 profiles for consider then rank, 16 versus 32 profiles for full rank) suggest an interesting managerial trade-off between predictive ability and task time. With specific loss functions on predictability and task time, such comparisons enable managers to design more efficient market research studies.

Aspects vs. Features

Finally, we address whether respondents process profiles by features or by aspects when they use lexicographic processes. Recall that lexicographic by features (LBF) is a restricted form of LBA where respondents rank by features (e.g., Verizon versus Sprint versus Nextel versus Cingular) rather than aspects (Verizon versus not Verizon). Because LBA nests LBF, LBA's *fit* statistics will be better. However, there is no guarantee that LBA's holdout predictions will be better than those of LBF. If respondents process profiles by features, then LBF could predict as well as LBA, perhaps better if LBA exploits random variations.

Table 1 compares LBA to LBF. On average, LBA predicts significantly better on both holdout violations and hit rates. LBA predicts better in all four cells and significantly better in three of the four cells (t's = 1.8, 7.1, 2.4, and 4.5; p's = 0.07, 0.00, 0.02, and 0.00 in cells 1–4). However, LBF predicts better for about a third of the respondents (35% for holdout violations and 34% for hit rates, with no significant differences between experimental cells).

7. Analysis of Computer Data from a Study by Lenk et al. (1996)

We were fortunate to obtain a classic conjoint-analysis data set in which respondents evaluated full profiles of computers that varied on 13 binary features: telephone service hotline, amount of memory, screen size, CPU speed, hard disk size, CD ROM, cache, color, availability, warranty, bundled software, guarantee, and price. Respondents were presented with 16 full

[22] This observation is tempered with the realization that the full-rank cells provide more ordered pairs than the consider-then-rank cells (496 versus 183, on average).

[23] For many respondents, the hit-rate prediction of LBA is tied with HBRL(4). Among those that are not tied, significantly more fit better with LBA in the full-rank cells than in the consider-then-rank cells ($t = 2.3$, $p = 0.02$).

[24] The comparisons are enjoyment, interest, and accuracy (2.07 versus 2.04, $t = 0.1$, $p = 0.90$); task time (3.40 versus 3.75 minutes, $t = 0.5$, $p = 0.64$) for 16 versus 32 profiles in a consider-then-rank task.

[25] Hit rates are worse by 2.9 percentage points, but the difference is not significant ($t = 1.7$, $p = 0.00$). Because the predicted holdout percentages are based only on the full-rank cells, they differ slightly from those in Table 1.

profiles and asked to provide a rating on a 10-point likelihood-of-purchase scale. They were then given a holdout task in which they evaluated four additional profiles on the same scale. These data were collected and analyzed by Lenk et al. (1996), who suggest excellent fit and predictive ability with hierarchical Bayes compensatory models. Based on their analysis and our intuition, we felt that the features in this study were more likely to be compensatory than those in the SmartPhone study. However, this is an empirical question.[26]

We first degraded the data from ratings to ranks. For example, if Profile A were rated as a 10 and Profile B were rated as a 1, we retained only that Profile A was preferred to Profile B. Because there were 10 scale points and 16 profiles, there were many ties—an average of 6.6 unique ratings per respondent. Interestingly, even though there were many ties, there were approximately 96 ranked pairs of profiles per respondent—80% of what would be obtained with full ranks. Because the degraded data are partial ranks, we can analyze the data with greedoid methods and compare predictions to HBRL(∞) and HBRL(q).[27]

Table 2 reports the fit and prediction results for the computer data. As with the SmartPhone data, we address the predictive ability of LBA compared to (1) an unconstrained additive model and (2) a q-compensatory model. On these data, the unconstrained additive model predicts better than LBA, significantly so for holdout pairs. (The difference in hit rates is only 1 respondent out of 201 respondents.) However, LBA predicts significantly better than the q-compensatory model. Comparisons for other values of q are available in a supplemental appendix available on the *Marketing Science* Web site.

For the computer data, LBA predicts better for 58% of the respondents compared to 25% for HBRL(4); the remainder are tied. We distinguish fewer respondents by hit rate because hit-rate classification is a coarser measure: 32% LBA, 20% HBRL(4), and 47% tied.

Interestingly, LBA on the degraded data does as well as *metric* hierarchical Bayes on the ratings data (0.687; Lenk et al. 1996, p. 181) and better than either OLS (0.637; Lenk et al. 1996, p. 181) and latent class analysis (0.408; Lenk et al. 1996, p. 181).[28] In this case, a reduction in effort (ranking versus rating) might have had little effect on predictive ability. For further discussion of ranking versus rating data, see Huber et al. (2002).

Table 2 Comparison of Fit and Prediction for Computer Data (Lenk et al. 1996)

	Lexicographic by aspects	Hierarchical Bayes ranked logit	Hierarchical Bayes ranked logit ($q = 4$)
Fit (percent pairs)	0.899*	0.906*	0.779
Holdout (percent pairs)	0.790*	0.827**	0.664
Holdout hit rate	0.686*	0.692*	0.552

*LBA and HBRL significantly better than HBRL(4)—at 0.05 level. **HBRL significantly better than LBA and HBRL(4).

Table 2 is consistent with the analysis of metric data by Jedidi and Kohli (2005), who found that a different lexicographic model (binary satisficing, LBS) fit almost as well as an unconstrained additive model (0.93 fit pairs for LBS versus 0.95 for classic LINMAP; no data available on holdouts). The Jedidi and Kohli (2005) context is remarkably similar to that of Lenk et al. (1996): metric ratings of 16 laptop computers described by memory, brand, CPU speed, hard drive size, and price (in a $3^3 2^2$ fractional design).

Comparing the SmartPhone and computer data, we get surprisingly similar respondent-level comparisons. LBA predicts at least as well as HBRL(4) for 57% of the SmartPhone respondents and 75% of the computer respondents.[29] Jedidi and Kohli (2005) did not test a q-compensatory model, but they did find that an unconstrained additive model was not significantly different from LBS for 67% of their respondents. Thus, on all data sets for more than half of the respondents, noncompensatory models predict holdout data at least as well as q-compensatory models.

We can also compare the predictive ability of LBA to an unconstrained additive model. LBA predicts at least as well as HBRL for 49% of the SmartPhone respondents and 62% of the computer respondents. Thus, even compared to an unconstrained additive model, LBA is promising as a predictive tool.

8. Managerial Implications

Manufacturers, retailers, or Web site designers seek to design products, store layouts, or Web sites that have (or emphasize) those aspects that strongly influence which products customers select for further consideration. They seek to avoid those aspects that customers use to eliminate products. In the parlance of product development, these are the "must-have" or "must-not-have" aspects or features (Hauser et al. 2005). Both General Motors and Nokia have indicated

[26] There are other differences between the data sets that are worth further study. For example, the rating task might induce more compensatory processing than the full-rank or consider-then-rank tasks.

[27] For the Lenk et al. (1996) data, HBRL predictions are significantly better than those by LINMAP. For holdout pairs, LINMAP predicts 0.734 ($t = 5.3$, $p = 0.00$). For hit rates, LINMAP predicts 0.597 ($t = 2.6$, $p = 0.01$).

[28] We compare to the highest hit rate they report—that for hierarchical Bayes estimated with 12 profiles. For 16 profiles, they report a hit rate of 0.670. For other statistics, hierarchical Bayes with 16 profiles performs better than with 12 profiles (Lenk et al. 1996, p. 181).

[29] The corresponding percentages for hit rates are 71% and 80%.

Table 3 Top Lexicographic Aspects for SmartPhones (for Our Sample)

Aspect	ABA or EBA	Affect consideration* (%)	Top aspect† (%)
Price—$499	EBA	49.2	26.1
Flip	ABA	32.0	10.4
Small	ABA	29.4	10.0
Price—$299	EBA	19.8	4.2
Keyboard	ABA	17.3	7.5
Price—$99	ABA	14.5	4.8

*Column sums to 300% over all aspects. †Column sums to 100% across all aspects. Most aspects not shown.

to us that the identification of must-have aspects is an extremely important goal of their product development efforts (private communication). Table 3 lists the six aspects that were used most often by SmartPhone respondents and indicates whether they were used to retain profiles as in ABA or eliminate profiles as in EBA (second column), the percentage of consumers who used that aspect as one of the first three aspects in a lexicographic order (third column), and the percent who used that aspect as the first aspect in a lexicographic order (fourth column).

Table 3 has a number of implications. First, for our student sample, there are clear price segments—for almost half the sample, high price is an EBA aspect. Second, "flip" and "small" are each ABA aspects for about 30% of the respondents. For this sample, any manufacturer would lose considerable market share if it did not include SmartPhones that were small and flip. The keyboard aspect is interesting. Keyboard is an ABA aspect for 17.3% of the respondents and an EBA aspect for 7.5% of the respondents (not shown). On this aspect, a manufacturer would be best advised to offer both SmartPhones with keyboards and SmartPhones without keyboards. Finally, brand, service provider, and operating system are not high in the summary of lexicographic orderings.

It is interesting that in our data price aspects were often, but not always, EBA aspects, while all other aspects were ABA aspects. (This is true for aspects not shown in Table 3.) We do not know if this generalizes to other categories. Furthermore, although "high price" was the top lexicographic aspect in our study, this could be a consequence of the category or our student sample. We do not expect price to be the top lexicographic aspect in all categories nor do we feel that this result affected the basic scientific and methodological findings about lexicographic processing or predictive ability.

9. Summary, Conclusions, and Future Research

In this paper, we propose methods to estimate noncompensatory process descriptions with either full-rank or partial-rank data. Estimation is a nontrivial combinatorial problem which has hitherto been too time-consuming to solve. Greedoids provide a structure and theory to transform an $n!$ problem into a 2^n problem, which for practical problems decreases running time by a factor the order of 10^{13}.

We tested greedoid methods empirically for SmartPhones and computers. The data suggest that the estimated lexicographic models predict well. Noncompensatory models predict at least as well as compensatory models for more than half of the respondents in both studies. Greedoid methods are flexible. We applied the methods with full-rank, consider-then-rank, and degraded ratings tasks, but the methods apply to any partial-order task, including repeated choice tasks. We believe they are promising for the study of noncompensatory decision rules.

Based on simulations, the SmartPhone data, and the computer data, we present the following summary.

Methodological

- It is feasible to estimate noncompensatory processes with a greedoid dynamic program.
- Noncompensatory estimates predict holdout pairs and holdout hit rates well.
- Greedoid methods appear to be robust—they predict well even though additive models can represent lexicographic processes.
- A consider-then-rank task reduces task time, increases completion rates, and improves perceived enjoyment, accuracy, and interest, although at some loss in predictive ability.
- Doubling the number of profiles from 16 to 32 improves predictive ability slightly for the full-rank task, but has no significant effect for the consider-then-rank task.
- Enabling respondents to sort profiles by aspects is seen as more difficult and time-consuming, but does not increase predictive ability.

Consumer Behavior

- Compared to a q-compensatory model, LBA predicts at least as well for more than half of the respondents.
- More respondents appear to process *aspects* rather than *features* lexicographically.
- The full-rank task (versus a consider-then-rank task) *does* increase the percent of respondents for whom a lexicographic model fits better than a q-compensatory model.
- Enabling respondents to sort profiles *does not* increase the percentage of respondents for whom LBA predicts better than a q-compensatory model.
- Increasing the number of profiles in a consider-then-rank task *does not* increase the percentage of respondents for whom LBA predicts better than a q-compensatory model; this varies from the literature

that finds that increasing the number of profiles *does* increase heuristic processing for a choice (rank) task.

Managerial (for Our Sample and Category)

- Price is often used as an EBA aspect. Nonprice aspects seem to be used as ABA aspects.
- "Small" and "flip" are key aspects for SmartPhones.

Future Directions

Greedoid methods provide a promising tool to study consumer behavior. Researchers can use the greedoid inference engine to investigate many impacts of consumers' constructive judgment and decision processes—manipulations that might be too intrusive if implemented by verbal protocols or information display tasks.

Methodologically, the exact dynamic program is still exponential in the number of aspects. We handled 16 aspects in 1.85 seconds. At this rate, greedoid methods can be used to evaluate up to 21 aspects in under a minute. However, we can handle much larger problems if we concentrate on the first few lexicographic aspects in a respondent's LBA process. Because the theory applies to partial orders, we can stop the dynamic program after m aspects yielding a running time proportional to ${}_nC_m$. For example, we could identify the top 5 out of 50 aspects in approximately 1 minute. Partial-order greedoid methods can be used to identify satisficing processes in which some aspect levels are considered as equivalent by respondents. Other heuristics might also be used (Kohli and Jedidi 2004, Kohli et al. 2006).

Finally, there are interesting commonalities and differences between the SmartPhone and computer data sets and, perhaps, some empirical generalizations.

Acknowledgments

This research was supported by Massachusetts Institute of Technology (MIT) Sloan School of Management, Center for Innovation in Product Development, and Operations Research Center. It was also supported by the Office of Naval Research Contract N00014-98-1-0317. This paper may be downloaded from http://mitsloan.mit.edu/vc. The authors offer special thanks to Peter Lenk, who kindly and unselfishly shared data with them. The authors also thank Ashvini Thammaiah, who assisted in the development and pretesting of the Web-based questionnaire. The study benefited from comments by numerous pretests at the MIT Operations Research Center and the Kappa Alpha Theta Sorority at MIT. SmartPhone images were produced by R. Blank. The authors thank Theodoros Evgeniou, Shane Frederick, Steven Gaskin, Rejeev Kohli, and Olivier Toubia for insightful comments on an earlier draft. This paper benefited from presentations at the MIT Operations Research Center; MIT Center for Product Development; MIT Marketing Group Seminar; Michigan's Ross School of Business Seminar Series; 2004 Marketing Science Conference in Rotterdam, The Netherlands; 2005 Marketing Science Conference at Emory University; AMA 2004 Advanced Research Techniques Forum, Whistler, British Columbia; 2004 EXPLOR Award Case Study Showcase, New Orleans, Louisiana; 2004 Management Roundtable Voice-of-the-Customer Conference, Boston, Massachusetts; 2006 Sawtooth Software Conference, Delray Beach, Florida; and the Marketing Camp at Columbia University, New York. The final three author names are listed alphabetically. Contributions were extensive and synergistic.

Appendix. Proofs of the Formal Propositions

Proposition 1. *Let E be a set of aspects, and let X be a partial order on the profiles. Let G be the collection of ordered subsets of E that are lexicoconsistent with X. Then, G is a greedoid language.*

Proof. We show that greedoid properties (2) and (3) hold for collection G. Property (1) is implied by (2). Property (2): Lexicoconsistent means that there is no pair of profiles P and P' with $P \succ_L P'$ and $P' \succ_X P$. So, if L is consistent with X, then $L\backslash e$ is consistent with X because the relations with respect to $L\backslash e$ are a subset of the relations with respect to L. Property (3): Let e be the first aspect in L' such that $e \notin L$. Such an e is guaranteed to exist since $|L'| > |L|$. We show that $(L, e) \in G$ via a contradiction. Suppose that there are profiles P and P' with $P \succ_{L,e} P'$ and $P' \succ_X P$. Because L and X are consistent, it follows that P and P' are unrelated with respect to L and, thus, $P \succ_e P'$. Let L'' be aspects in L' prior to e. Then, $L'' \subseteq L$, so P and P' are unrelated in L''. It follows that $P \succ_{L'',e} P'$ and, thus, $P \succ_{L'} P'$, contradicting that L' is consistent with X. We conclude that Property 3 is true.

Proposition 2. *Let L and L' be two different permutations of a subset E' of aspects, and let e be any aspect not in E'. Then, the number of inconsistencies directly caused by e in (L, e) is the same as the number of inconsistencies caused by e in (L', e).*

Proof. Suppose that for profiles P and P', $P \succ_X P'$. Aspect e causes an inconsistency with respect to profiles P and P' in (L, e) if and only if the following conditions hold: (i) profiles P and P' are undifferentiated by the aspects in L, and (ii) $P' \succ_e P$. These are the same conditions under which e causes an inconsistency with respect to P and P' in (L', e).

Proposition 3. *If weights are associated with each ordered pair in X, then* (1) *the new G is a greedoid language*, (2) *Algorithm* 1 *determines whether there exists an L consistent with X*, (3) *Proposition* 2 *extends to the new G, and* (4) *Algorithm* 2 *finds the best lexicographic description if $c(\cdot)$ is redefined to $c(S\backslash\{i\}, i) =$ sum of weights of violations caused by aspect i following set $S\backslash\{i\}$.*

Proof. Proposition 1 and Algorithm 1 are unaffected by nonunit weights because the determination of consistency does not depend on the weights associated with inconsistencies. Proposition 2 extends easily to the case where nonunit weights are allowed. With essentially the same proof, it can be shown that the sum of the weights of inconsistencies directly caused by aspect e in order (L, e) is still independent of the permutation of the preceding aspects L. With the redefinition of $c(S\backslash\{i\}, i)$, the validity of the new dynamic programming formulation follows from the extension of Proposition 2.

Proposition 4. *Suppose that L is an ordering of a subset of aspects that is consistent with the preferences of X. Then, there is an optimal ordering of aspects that begins with the order L.*

Proof. Suppose that L' is an optimal ordering of aspects, that is, it is the one that minimizes the number of inconsistencies with respect to X. Let L'' be obtained from L' by moving the aspects of L to the front of the order. We will show that any inconsistency with respect to L'' is also an inconsistency with respect to L', thus showing that L'' is at least as optimal as L'. Suppose that for profiles P and P', $P \succ_X P'$ and $P' \succ_{L''} P$. Let e be the first aspect in L'' that differentiates P and P'. Because L is consistent with X, it follows that $e \notin L$. Therefore, e is also the first aspect in L' that differentiates P and P', so $P' \succ_{L'} P$. It follows that P and P' also cause an inconsistency with respect to L', proving the proposition.

References

Bettman, J. R., L. W. Park. 1980. Effects of prior knowledge and experience and phase of the choice process on consumer decision processes: A protocol analysis. *J. Consumer Res.* **7**(3) 234–248.

Bettman, J. R., M. F. Luce, J. W. Payne. 1998. Constructive consumer choice processes. *J. Consumer Res.* **25**(3, Dec) 187–217.

Björner, A., G. M. Ziegler. 1992. Introduction to greedoids. N. White, ed. *Matroid Applications, Encyclopedia of Mathematics and Its Applications*, Vol. 40. Cambridge University Press, London, UK, 284–357.

Bröder, A. 2000. Assessing the empirical validity of the "Take the Best" heuristic as a model of human probabilistic inference. *J. Experiment. Psych.: Learn., Memory, Cognition* **26**(5) 1332–1346.

Cui, D., D. Curry. 2005. Prediction in marketing using the support vector machine. *Marketing Sci.* **24**(4), 595–615.

Dawes, R. M. 1979. The robust beauty of improper linear models in decision making. *Amer. Psychologist* **34** 571–582.

DeSarbo, W. S., D. R. Lehmann, G. Carpenter, I. Sinha. 1996. A stochastic multidimensional unfolding approach for representing phased decision outcomes. *Psychometrica* **61**(Sept) 485–508.

Einhorn, H. J. 1970. The use of nonlinear, noncompensatory models in decision making. *Psych. Bull.* **73**(3) 221–230.

Einhorn, H. J., R. M. Hogarth. 1981. Behavioral decision theory: Processes of judgment and choice. *Annual Rev. Psych.* **32** 52–88.

Evgeniou, T., C. Boussios, G. Zacharia. 2005. Generalized robust conjoint estimation. *Marketing Sci.* **24**(3) 415–429.

Fishburn, P. C. 1974. Lexicographic orders, utilities and decision rules: A survey. *Management Sci.* **20**(11) 1442–1471.

Frederick, S. 2005. Cognitive reflection and decision making. *J. Econom. Perspectives* **19**(2) 25–42.

Gensch, D. H. 1987. A two-stage disaggregate attribute choice model. *Marketing Sci.* **6**(Summer) 223–231.

Gensch, D. H., E. S. Soofi. 1995. Information-theoretic estimation of individual consideration sets. *Internat. J. Res. Marketing* **12**(May) 25–38.

Gigerenzer, G., D. G. Goldstein. 1996. Reasoning the fast and frugal way: Models of bounded rationality. *Psych. Rev.* **1003**(4) 650–669.

Gilbride, T., G. M. Allenby. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Sci.* **23**(3, Summer) 391–406.

Green, P. E., A. M. Krieger, P. Bansal. 1988. Completely unacceptable levels in conjoint analysis: A cautionary note. *J. Marketing Res.* **25**(Aug) 293–300.

Griffin, A., J. R. Hauser. 1993. The voice of the customer. *Marketing Sci.* **12**(1, Winter) 1–27.

Hauser, J. R. 1978. Testing the accuracy, usefulness and significance of probabilistic models: An information theoretic approach. *Oper. Res.* **26**(3, May–June) 406–421.

Hauser, J. R., O. Toubia. 2006. The impact of utility balance and endogeneity in conjoint analysis. *Marketing Sci.* **24**(3) 498–507.

Hauser, J. R., B. Wernerfelt. 1990. An evaluation cost model of consideration sets. *J. Consumer Res.* **16**(Mar) 393–408.

Hauser, J. R., G. Tellis, A. Griffin. 2005. Research on innovation: A review and agenda for marketing science. *Marketing Sci.* **25**(6) 687–717.

Held, M., R. M. Karp. 1962. A dynamic programming approach to sequencing problems. *SIAM J. Appl. Math.* **10**(1) 196–210.

Huber, J., D. Ariely, G. Fischer. 2002. Expressing preferences in a principal-agent task: A comparison of choice, rating and matching. *Organ. Behav. Human Decision Processes* **87**(1, Jan) 66–90.

Jedidi, K., R. Kohli. 2005. Probabilistic subset-conjunctive models for heterogeneous consumers. *J. Marketing Res.* **42**(3) 483–494.

Jedidi, K., R. Kohli, W. S. DeSarbo. 1996. Consideration sets in conjoint analysis. *J. Marketing Res.* **33**(Aug) 364–372.

Johnson, E. J., R. J. Meyer. 1984. Compensatory choice models of noncompensatory processes: The effect of varying context. *J. Consumer Res.* **11**(1, June) 528–541.

Johnson, E. J., R. J. Meyer, S. Ghose. 1989. When choice models fail: Compensatory models in negatively correlated environments. *J. Marketing Res.* **26**(Aug) 255–290.

Johnson, R. 1991. Comment on adaptive conjoint analysis: Some caveats and suggestions. *J. Marketing Res.* **28**(May) 223–225.

Kendall, M. G. 1975. *Rank Correlation Methods*. Charles Griffin & Company, Ltd., London, UK.

Kim, J. G. 2004. Dynamic heterogeneous choice heuristics: A Bayesian hidden Markov mixture model approach. Working paper, MIT Sloan School of Management, Cambridge, MA.

Klein, N. M. 1988. Assessing unacceptable attribute levels in conjoint analysis. *Adv. Consumer Res.* **14** 154–158.

Kohli, R., K. Jedidi. 2004. Representation and inference of lexicographic preference models and their variants. *Marketing Sci.* **26**(3) 380–399.

Kohli, R., R. Krishnamurthi, K. Jedidi. 2006. Subset-conjunctive rules for breast cancer diagnosis. *Discrete Appls. Math.* **154**(7) 1100–1132.

Korte, B., L. Lovász. 1985. Basis graphs of greedoid and two-connectivity. *Math. Programming Stud.* **24** 158–165.

Korte, B., L. Lovász, R. Schrader. 1991. Greedoids. *Algorithms and Combinatorics Series*, Vol. 4. Springer-Verlag, Berlin, Germany.

Lenk, P. J., W. S. DeSarbo, P. E. Green, M. R. Young. 1996. Hierarchical Bayes conjoint analysis: Recovery of partworth heterogeneity from reduced experimental designs. *Marketing Sci.* **15**(2) 173–191.

Liechty, J. C., D. K. H. Fong, W. S. DeSarbo. 2005. Dynamic models incorporating individual heterogeneity: Utility evolution in conjoint analysis. *Marketing Sci.* **24**(3) 285–293.

Lohse, G. J., E. J. Johnson. 1996. A comparison of two process tracing methods for choice tasks. *Organ. Behav. Human Decision Processes* **68**(1, Oct) 28–43.

Luce, M. F., J. W. Payne, J. R. Bettman. 1999. Emotional trade-off difficulty and choice. *J. Marketing Res.* **36** 143–159.

Malhotra, N. 1986. An approach to the measurement of consumer preferences using limited information. *J. Marketing Res.* **23**(Feb) 33–40.

Martignon, L., U. Hoffrage. 2002. Fast, frugal, and fit: Simple heuristics for paired comparisons. *Theory Decision* **52** 29–71.

Mitchell, T. M. 1997. *Machine Learning*. WCB McGraw-Hill, Boston, MA.

Montgomery, H., O. Svenson. 1976. On decision rules and information processing strategies for choices among multiattribute alternatives. *Scandinavian J. Psych.* **17** 283–291.

Nakamura, Y. 2002. Lexicographic quasilinear utility. *J. Math. Econom.* **37** 157–178.

Niño-Mora, J. 2000. On certain greedoid polyhedra, partially indexable scheduling problems, and extended restless bandit allocation indices. Economics Working Papers 456, Department of Economics and Business, Universitat Pompeu Fabra.

Olshavsky, R. W., F. Acito. 1980. An information processing probe into conjoint analysis. *Decision Sci.* **11**(July) 451–470.

Oppewal, H., J. J. Louviere, H. J. P. Timmermans. 1994. Modeling hierarchical conjoint processes with integrated choice experiments. *J. Marketing Res.* **31**(Feb) 92–105.

Payne, J. W. 1976. Task complexity and contingent processing in decision making: An information search. *Organ. Behav. Human Performance* **16** 366–387.

Payne, J. W., J. R. Bettman, E. J. Johnson. 1988. Adaptive strategy selection in decision making. *J. Experiment. Psych.: Lear., Memory, Cognition* **14** 534–552.

Payne, J. W., J. R. Bettman, E. J. Johnson. 1993. *The Adaptive Decision Maker*. Cambridge University Press, Cambridge, UK.

Reeves, R. 1961. *Reality in Advertising*. Knopf Publishing, New York.

Roberts, J. H., J. M. Lattin. 1991. Development and testing of a model of consideration set composition. *J. Marketing Res.* **28**(Nov) 429–440.

Rossi, P. E., G. M. Allenby. 2003. Bayesian statistics and marketing. *Marketing Sci.* **22**(3, Summer) 304–328.

Shugan, S. M. 1980. The cost of thinking. *J. Consumer Res.* **7**(2, Sept) 99–111.

Srinivasan, V. 1998. A strict paired comparison linear programming approach to nonmetric conjoint analysis. J. E. Aronson, S. Zionts, eds. *Operations Research: Methods, Models, and Applications*. Quorum Books, Westport, CT, 97–111.

Srinivasan, V., C. S. Park. 1997. Surprising robustness of the self-explicated approach to customer preference structure measurement. *J. Marketing Res.* **34**(May) 286–291.

Srinivasan, V., A. Shocker. 1973. Linear programming techniques for multidimensional analysis of preferences. *Psychometrika* **38**(3, Sept) 337–369.

Srinivasan, V., G. A. Wyner. 1988. Casemap: Computer-assisted self-explication of multiattributed preferences. W. Henry, M. Menasco, K. Takada, eds. *Handbook on New Product Development and Testing*. D. C. Heath, Lexington, MA, 91–112.

Swait, J. 2001. A noncompensatory choice model incorporating cutoffs. *Transportation Res.* **35**(Part B) 903–928.

Thorngate, W. 1980. Efficient decision heuristics. *Behavioral Sci.* **25**(May) 219–225.

Toubia, O., D. I. Simester, J. R. Hauser, E. Dahan. 2003. Fast polyhedral adaptive conjoint estimation. *Marketing Sci.* **22**(3, Summer) 273–303.

Tversky, A. 1972. Elimination by aspects: A theory of choice. *Psych. Rev.* **79** 281–299.

Urban, G. L., J. R. Hauser. 2004. "Listening-in" to find and explore new combinations of customer needs. *J. Marketing* **68**(Apr) 72–87.

Zaltman, G. 1997. Rethinking market research: Putting people back in. *J. Marketing Res.* **23**(Nov) 424–437.