Marketing Papers

Wharton Faculty Research

12-2008

Sequential Sampling Models of Choice: Some Recent Advances

Thomas Otter

Joe Johnson

Jörg Rieskamp

Greg M. Allenby

Jeff D. Brazell

See next page for additional authors

Sequential Sampling Models of Choice: Some Recent Advances

Abstract

Choice models in marketing and economics are generally derived without specifying the underlying cognitive process of decision making. This approach has been successfully used to predict choice behavior. However, it has not much to say about such aspects of decision making as deliberation, attention, conflict, and cognitive limitations and how these influence choices. In contrast, sequential sampling models developed in cognitive psychology explain observed choices based on assumptions about cognitive processes that return the observed choice as the terminal state. We illustrate three advantages of this perspective. First, making explicit assumptions about underlying cognitive processes results in measures of deliberation, attention, conflict, and cognitive limitation. Second, the mathematical representations of underlying cognitive processes imply well documented departures from Luce's Choice Axiom such as the similarity, compromise, and attraction effects. Third, the process perspective predicts response time and thus allows for inference based on observed choices and response times. Finally, we briefly discuss the relationship between these cognitive models and rules for statistically optimal decisions in sequential designs.

Disciplines

Business | Marketing

Author(s)

Thomas Otter, Joe Johnson, Jörg Rieskamp, Greg M. Allenby, Jeff D. Brazell, Adele Diederich, Wesley Hutchinson, Steven MacEachern, Shiling Ruan, and Jim Townsend

Sequential Sampling Models of Choice: Some Recent Advances

by

Thomas Otter
J.W. Goethe Universität (Marketing)
otter@marketing.uni-frankfurt.de

Jörg Rieskamp
Max Planck Institute for Human Development
(Psychology)
rieskamp@mpib-berlin.mpg.de

Jeff D. Brazell
The Modellers, LLC (Marketing)
Jeff.Brazell@themodellers.com

Wes Hutchinson
University of Pennsylvania (Marketing)
jwhutch@wharton.upenn.edu

Shiling Ruan
Ohio State University (Statistics)
ruan@stat.osu.edu

Joe Johnson
Miami University (Psychology)
johnsojg@muohio.edu

Greg M. Allenby
Ohio State University (Marketing)
allenby.1@osu.edu

Adele Diederich
Jacobs University Bremen (Psychology)
a.diederich@iacobs-university.de

Steven MacEachern
Ohio State University (Statistics)
snm@stat.osu.edu

Jim Townsend
Indiana University (Psychology)
jtownsen@indiana.edu

January, 2008

Sequential Sampling Models of Choice: Some Recent Advances

January, 2008

Abstract

Choice models in marketing and economics are generally derived without specifying the underlying cognitive process of decision making. This approach has been successfully used to predict choice behavior. However, it has not much to say about such aspects of decision making as deliberation, attention, conflict, and cognitive limitations and how these influence choices. In contrast, sequential sampling models developed in cognitive psychology explain observed choices based on assumptions about cognitive processes that return the observed choice as the terminal state. We illustrate three advantages of this perspective. First, making explicit assumptions about underlying cognitive processes results in measures of deliberation, attention, conflict, and cognitive limitation. Second, the mathematical representations of underlying cognitive processes imply well documented departures from Luce's Choice Axiom such as the similarity, compromise, and attraction effects. Third, the process perspective predicts response time and thus allows for inference based on observed choices and response times. Finally, we briefly discuss the relationship between these cognitive models and rules for statistically optimal decisions in sequential designs.

Sequential Sampling Models of Choice: Some Recent Advances

1. Introduction

Choice models in marketing and economics are usually derived using a constrained maximization framework. The decision maker is assumed to choose the option that maximizes utility subject to a budget constraint. Observed departures from systematic behavior as implied by a particular utility function are attributed to aspects of utility that cannot be observed by the analyst. This viewpoint is generally not distinguishable from the assumption that some aspect of decision making is truly random. While the prevalent constrained maximization framework could in principle be extended to accommodate any number and type of constraints such as constraints related to attention, cognitive capacity, the time to make the decision, etc. such an extension has proven to be difficult (see Gilbride & Allenby 2006 for an example that derives optimal screening rules). Consequently, the prevailing choice models used in marketing and economics, such as the logit and the probit model, are consistent with instantaneous utility maximization, i.e. they abstract away from the cognitive processes that lead to the identification of the alternative chosen from a set.

Adamovicz et al. (this volume) provide a general discussion of ways to improve on currently used choice models. In this paper, we focus on sequential sampling models derived from primitive assumptions about the cognitive processes that result in a choice. The basic idea is that the observed choice corresponds to the terminal state of a process that started in some state and evolved over a period of time. This is obviously closer to reality than is instantaneous utility maximization. However, given that the cognitive process itself is unobservable, the crucial question is whether the description of the underlying cognitive process leads to different inferences in comparison to standard models. In a sense, we are therefore documenting how

2

integrating with respect to the realization of a cognitive process results in predictions that are more in line with the observed data than integrating with respect to the commonly assumed distributions for the error term in random utility models.

We do not claim that a process perspective is necessary to achieve these results. In theory, every process corresponds to some (non-standard) error distribution. We also do not claim that a process perspective is fundamentally at odds with a random utility framework since one may choose to refer to realizations of the process as the random components of utility and to non-stochastic elements in the process description as the 'deterministic' component of utility. However, we hope to illustrate that a process perspective is a natural starting point for the construction of choice models that meaningfully extend the set of models currently in use. We also show how a process perspective leads to measures for other aspects of choice such as response time, diligence, amount of conflict and processing capacity. Such measures are useful for applications that go beyond the mere prediction of choice outcomes. Finally, we will introduce a very recent stream of research that compares the performance implications of process models to statistically optimal decisions derived from sequential probability ratio testing (Bogacz et al. 2006). This line of research is interesting because it elevates (some) process models beyond the status of useful descriptions by motivating their architecture from optimal decision making under time constraints given noisy input information.

The remainder of the paper is organized as follows: In Section 2 we introduce sequential sampling models that motivate choice through latent (cognitive) accumulation processes and briefly review their history. Section 3 discusses selected applications. Section 4 illustrates some challenges for estimation and Section 5 concludes with a discussion and directions for future research.

2. Sequential Sampling Models

Sequential sampling models (e.g. Townsend & Ashby 1983) were originally developed in the context of simple perceptual identification tasks that require only low levels of cognition. The idea is that, after the onset of a stimulus, the decision maker sequentially extracts and accumulates information from the stimulus and/or its mental representation to determine the nature of the stimulus. In the simplest task, the decision maker has to determine if a stimulus is present or not. Another popular task consists of pairs of letters where the decision maker has to identify pairs of same or different letters. In this context, the decision maker's perceptual and cognitive system is summarized by a rate of information accumulation and a decision criterion that specifies the amount of information required before a response occurs. Together the accumulation rate and the decision criterion determine response probabilities and response times.

Race models, or counter and accumulator models, assume that evidence in favor of the available options is accumulated in separate stores (e.g., LaBerge 1962, Vickers 1970). The accumulation process stops whenever any of the stores first accumulates a prespecified amount of evidence, and the corresponding choice is made. The time at which this happens corresponds to the response time. In applications, race models often build on the assumption that evidence accrues to the stores in the form of discrete 'hits' that follow independent Poisson processes.

Diffusion models, or the related random walk models, assume that *relative* evidence is accumulated over time (e.g., Ashby 1983, Stone 1960). In the special case of two alternatives, relative evidence is defined as the evidence difference or the natural log of the ratio. With two alternatives, both definitions result in a scalar that is assumed to evolve according to a continuous stochastic process. Evidence for one alternative is therefore simultaneously evidence

4

against the other alternative. Relative evidence has to exceed / fall below prespecified boundaries for a response to occur. The particular boundary reached first determines the response, and the time at which the boundary is reached, the response time. Stochastic processes that have been studied in some detail in this context are the Wiener and the Orenstein-Uhlenbeck process (Diederich & Busemeyer 2003, Smith 2000). In practice it is common to approximate the continuous time stochastic processes by discrete time Markov chains (cf. Smith 2000).

Both race and diffusion models have successfully been applied to perceptual identification tasks. In more recent work they have been generalized to preferential choice among multi-attribute alternatives providing explanations for empirically well documented violations of Luce's choice axiom (see Rieskamp et al. 2006 for a recent overview).

Luce's choice axiom implies independence from irrelevant alternatives and regularity. Independence from irrelevant alternatives (IIA) requires that ratios of choice probabilities for fixed alternatives are constant across varying choice sets. A weaker version of IIA, order independence, only requires the order of probabilities to be constant across varying choice sets (Tversky 1972b). Regularity requires that adding alternatives to a set always translates into weakly smaller probabilities of choosing any of the original options.

The *similarity* and the *compromise* effects both violate IIA. The *attraction* effect (Huber, Payne & Puto 1982, Huber & Puto 1983) refutes regularity. The *similarity effect* (Tversky 1972a), in the limit, refers to the situation where the addition of another alternative to a set decreases the probability of choosing alternatives perceived to be identical but leaves choice probabilities of other alternatives unchanged. The *compromise effect* (Simonson 1989, Tversky & Simonson 1993) describes the observation that a compromise, alternative C is most likely chosen from the set {A, B, C}, even though the decision maker does not show a preference for C

in either of the binary choices between {A, C} and {B, C}. The compromise option C is characterized by attribute levels that fall in between the attribute values of options A and B, hence the name. Finally, the *attraction effect* describes the observation that adding an asymmetrically dominated alternative to a set of two alternatives increases the choice probability of the now respectively dominant alternative.

Busemeyer & Townsend (1993) introduced decision field theory (DFT) as a framework for process based modeling of choice. DFT is built on the idea that relative evidence has to exceed / fall below prespecified boundaries for a choice to occur. Thus, it can be viewed as an instance of a diffusion model. Roe et al. (2001) generalized the model to preferential choices among multi-attribute alternatives. DFT contains classic random utility theory as a special case and provides a unifying framework for explaining the described violations of Luce's choice axiom.

DFT assumes that the decision maker's attention fluctuates stochastically between the various attributes. At any given moment, a single attribute is considered and options with similar attribute values accumulate similar amounts of evidence. Plotting accumulated evidence against time, the accumulation paths of similar options thus exhibit positive correlation. Similar alternatives therefore tend toward meeting the decision criterion simultaneously, creating the similarity effect.

DFT further assumes that the accumulated evidence in favor of alternatives is subject to decay and competition. Without any processing input, evidence in favor of alternatives gradually decays to a state of indifference. Competition is incorporated in the model via inhibitory links between options that imply that increasing evidence for one option causes evidence for other options connected via inhibitory links to it to decrease over time. The

6

parameters quantifying the strength of competitive inhibition are a function of the distance between alternatives in the attribute space. Close alternatives inhibit each other more strongly than do distant alternatives.

The inhibitory links produce both the compromise and attraction effects. The inhibitory connection between the compromise option 'in the middle' and each of the extreme options is stronger than that between the extreme options. This asymmetry induces positive correlation between the relative evaluations of the extreme options and thus makes the choice of the compromise option more likely (cf. Kivetz et al. 2004, Usher & McClelland 2004).

In case of the attraction effect, the dominated alternative accumulates negative relative evidence, which through the inhibitory, negative link translates into a boost for the similar, but dominating option. Moreover, the dynamics of decision field theory link the size of these effects to the amount of time spent with the decision. The amount of time invested is in turn a function of the accuracy goal of the decision maker. For instance, the effect of inhibition accumulates over time, with the implication that the attraction effect only occurs in well deliberated choices.

DFT and its various extensions (see Section 3) provide a rich framework for modeling preferential choice among multi-attribute alternatives. However, specification issues and computational challenges must be addressed for likelihood based inference to become practically feasible with this framework (see Section 4). This may be one reason why early applications of sequential sampling models in marketing, discussed in the following section, rely on Poisson race models and not decision field theory.

3. Selected Applications in Psychology and Marketing

Applications of Race models in marketing

Otter et al. (forthcoming) explored the idea that respondent diligence can be distinguished from respondents' tastes using a Poisson race model applied to conjoint data. In the Poisson race model, evidence in favor of an alternative is assumed to accrue according to a Poisson process of discrete units that are called 'hits'. The decision maker tracks the number of hits in favor of each alternative on specific counters. As soon as any one counter reaches its threshold, the corresponding alternative is chosen and the race terminates. The Poisson race model contains the multinomial logit model as a special case with a threshold equal to one.

Respondent diligence is measured by the threshold parameter. Thresholds larger than one give rise to choice probabilities that structurally depart from IIA such that the chances of choosing bad alternatives decrease disproportionally (see Ruan 2007 for a detailed mathematical analysis). And, it can be shown that, given the same amount of data, a larger threshold translates into more likelihood information about taste parameters or part-worths than does a smaller threshold.

The Poisson race model implies a joint density for choices and response times. Otter et al. found that the integration of response times requires the modeling of constrained (processing) capacity and heterogeneous processing speeds that change over the course of the task as a function of process priming. Their empirical results support endogeneity of response times as implied by the model. Quick response times point to easy decisions where at least one of the alternatives is outstanding, and slow response times point to hard decisions where the alternatives are less or equally attractive. The integration of response times marginally improves predictions.

Ruan et al. (forthcoming) introduced dependence between the alternative specific Poisson counters to model choices among pairs of credit cards in a conjoint experiment. The statistical

model is motivated from independent psychological processes underlying the valuation of individual attribute levels that have stochastic components. Realized valuations of attribute levels are integrated deterministically to the overall evidence in favor of a particular alternative. The alternative that first accrues an amount of evidence equal to the threshold required for a decision is chosen. The model thus combines elements of attribute- and alternative-based decision making strategies.

Dependence between alternative specific attribute counters is a direct consequence of shared realizations of stochastic valuations of attribute levels. Two alternatives that are close in the space set up by the attributes thus not only achieve similar expected overall evaluations but each pair of realized overall evaluations is similar. For two identical alternatives, any pair of realized overall evaluations is identical.

Ruan et al. illustrate how their model naturally handles attribute based dominance as a special case of similarity. Attribute based dominance occurs if one alternative in a set is strictly better than another on at least one attribute and not worse on all other attributes. In their model, two alternatives always share the realized valuation of worse attribute levels. Thus, the race between the dominating and the dominated alternative can only end with the dominating alternative winning or in a tie. They develop a tie-breaking rule that puts positive probability on choosing the dominated alternative to accommodate errors and / or satisficing behavior.

While the special case of dominance could be addressed in a variety of ways, Ruan et al. found that their model improves predictions considerably over the standard logit formulation, even in the case where the design does not feature dominance relationships. It seems that their model, which is consistent with psychological processes at the attribute evaluation level and thus exhibits similarity effects, is one step closer to a structural representation of the choice process.

9

Huang and Hutchinson (forthcoming) applied a version of the Poisson race model to jointly model three dependent variables, namely choices, reaction times and confidence ratings in a belief verification task. They show that the joint information contained in the three dependent variables as summarized by the model parameters improves predictions of a theoretically related dependent variable (i.e., attitudes) compared to direct use of choices, reaction times or confidence ratings. However, they also report that, for their data, reactions times add very little incremental information beyond what is contained in choices and confidence ratings, despite the fact that reaction times are affected by their experimental manipulations and are reasonably well-predicted by the Poisson race model. The substantive goal of their research was to compare belief verification to retrospective thought-listing as a measure of cognitive responses to persuasive communications. The Poisson model was used to detect the effects of specific thoughts during exposure to an advertisement on subsequent beliefs about the advertised product. Their experiments illustrate the effectiveness of estimated model parameters in predicting consumers' attitudes and the fact that these model-based measures can outperform traditional thought-listing when people are unwilling or unable to report certain thoughts.

Applications and Extensions of Decision Field Theory

Diederich (1997) extended DFT to multi-attribute decision field theory (MDFT). MDTF assumes that the preference process (technically, the diffusion process) has a specific input valence for each attribute. At any particular time during deliberation the decision maker's attention process may be operating on the process for one attribute and, during the next moment, attention either continues to operate on the process for that attribute or attention switches to

10

operate on the process for another attribute. Thus, MDTF handles serial processing of attributes as a special, degenerate case of state-dependent attention switching (cf. Roe et al. 2001).

MDFT accounts for several empirical findings in the context of decision making under time pressure. For instance, changes in the choice probabilities and more generally preference orders as a function of time constraints are viewed as the result of changes in the decision criterion, i.e. the amount of relative evidence necessary for a choice to occur. Under time pressure, the decision maker has to base the decision on less relative evidence that is likely produced before all attributes are considered or individual attributes can be reconsidered (for details see Diederich 2003a).

Diederich (2003b) showed how MDFT can be used to measure the amount of conflict induced by the desirability/undesirability of attribute values and the variability of outcomes. Diederich & Busemeyer (1999) related violations of stochastic dominance to conflict caused by negatively correlated payoffs. MDFT predicts this violation. The decision maker's attention changes from moment to moment, switching back and forth from one uncertain state to another during deliberation. While attending momentarily to a particular state, the decision maker compares consequences produced by each action under that state. These momentary comparisons are integrated over time to form an integrated preference for each action. When the payoffs are negatively correlated, the comparisons change their sign back and forth from positive to negative as attention fluctuates, producing up-and-down vacillations in preference that lead to violations of stochastic dominance. When payoffs are positively correlated, the comparison always produces a positive (or zero) increment favoring Action A over Action B, independent of the state to which the decision maker attends. In this case, preference for the dominant action always increases over time, so that stochastic dominance is satisfied.

11

Johnson & Busemeyer (2005, 2007) applied the DFT framework to tasks with alternative response types, as well as to cognitive subprocesses assumed to drive discrete choice. First, they developed a model that provides a way of mapping a single option onto a numeric value, which takes the form of a series of comparisons using the binary DFT choice model. Second, they formalized the decision weighting (e.g., probability weighting) process as one of sequential sampling that determines the sequence of attentional foci in MDFT.

The pricing extension to DFT (Johnson & Busemeyer, 2005) predicts the buying price, selling price, and certainty equivalent for an option, in addition to choice probabilities when the option is paired against another. It predicts the empirically documented reversals between choices and prices as well as reversals between buying and selling prices. The model assumes that a price is generated for an option by conducting a series of implicit comparisons between the option and candidate response prices until a suitable price is identified. That is, a set of candidate prices is defined, as is a probability distribution over this set to determine which price is considered first.

The first price considered is then compared to the option using the DFT choice process. If the candidate price is preferred, the price is decremented to produce the next candidate price, whereas if the option is preferred, the price is incremented to produce the next candidate price. With some probability, the comparison results in an indifference response. This indifference response occurs with a probability governed by an exit rate parameter, whenever the diffusion process is in the neutral or zero state. With this specification, the model predicts an entire distribution of response prices for any given option. The difference between pricing types (buying vs. selling), for instance, is captured by differences in the initial distribution – lower prices tend to be considered first for buying situations to minimize expenditure; higher prices are

12

more likely considered first for selling prices to maximize revenue. Other factors that could be modeled through the initial price distribution are reference prices or anchors set by competitive prices.

4. Inference

Challenges for likelihood-based inference

The practical feasibility of likelihood based inference is almost required for models to be useful in marketing because of data structures that necessitate the use of hierarchical models and the need to evaluate (managerial) loss functions for decision making. Likelihood based inference is relatively straightforward for the Poisson race models discussed in this paper, as the latent Poisson processes are assumed to be time homogenous with independent increments (across time) after possibly rescaling time. Thus, the unobserved hits can be integrated out analytically to yield closed form expressions for both the marginal likelihood of choices and the joint likelihood of choices, response times, and confidence.

The situation seems to be more complicated with DFT and its generalizations. A central concept of DFT is that attention fluctuates momentarily between attributes (Busemeyer & Townsend 1993). These fluctuations give rise to correlated preferences such that the strength of the correlation between two alternatives is a function of their distance in the attribute space. Another central concept is decay in the accumulated preferences coupled with negative feedback, i.e. inhibition between alternative specific preference stores (Roe et al. 2001). The inhibitory strength is again proportional to how similar or close two alternatives are in the attribute space. Inhibition between alternatives is essential to produce the attraction and the compromise effect. However, the joint identification status of the probability law governing attention fluctuation, of

13

attribute weights, and inhibition parameters has yet to be investigated in detail. The same holds for the joint identification of attribute weights and Diederich's (1997) generalizations of the probability law governing attention fluctuation in the absence of inhibition.

The joint density of choices and response times in decision field theory is the solution to a system of stochastic differential equations describing an, in the multi-alternative case vector-valued, inhomogenous diffusion process. Diederich & Busemeyer (2003) show how to compute marginal, conditional choice probabilities and response time distributions accurately using a discrete time Markov approximation (cf. Smith 2000). Transition probabilities in the discrete representation of the state space are derived from the parameters of the underlying continuous process; straightforward matrix algebra then yields marginal, conditional choice probabilities and response time distributions. However, with more than three alternatives to choose from, an explicit approximation to the state space becomes tedious, and suitable algorithms have yet to be developed. A potential alternative to the explicit representation of the state space is to augment the unobserved accumulation paths in the context of simulation based Bayesian inference via MCMC (Tanner & Wong 1987).

Model free inference about basic architectural features and capacity

The topic of model identification has received considerable attention in mathematical psychology under the title 'model mimicry'. It has been shown that models with radically different assumptions about the processing architecture such as parallel versus serial processing are not easily distinguished empirically (Townsend 1972, Townsend & Ashby 1983).

However, strong methods for assessing such strategic mechanisms in elementary cognitive action have been worked out in psychology and cognitive science (e.g., Townsend &

14

Ashby 1983, Townsend & Wenger 2004, Schweickert & Townsend 1989): 1. Architecture (parallel vs. serial processing vs. more complex networks). 2. Work-load capacity (e.g., how does performance change as a function of the number of attributes or when the number of attribute levels is increased?). 3. Decisional stopping rules (e.g., when should or does a process cease during information processing or attribute weighting?). 4. Stochastic independence vs. negative or positive dependence among attributes or choice objects. These methodologies are theory driven and specified in rigorous mathematical form.

A particularly successful example of these methodologies is the Double Factorial Paradigm, which delivers direct identification of items (1) through (3) and indirect evidence of (4). Ashby & Townsend (1986) developed other experimental designs which allow direct adjudication of the independence issue. An especially provocative finding using this paradigm is its ability to uncover not only the typical limited capacity nature of human information processing but situations involving perceptual unification (aka configural or Gestalt processing) where efficiency actually improves with increased workload (e.g., Townsend & Nozawa, 1995).

5. Discussion and Future Research

It seems clear to the authors that an individual's choices must be based on underlying cognitive processes. Modeling in marketing, however, has traditionally taken a top-down approach. Thus, the recognition of inconsistency of an individual's choices has led from deterministic utility to Luce's choice axiom and random utility; the recognition of the similarity effect has led from the multinomial logit model to the nested multinomial logit model; the inability of these models to predict choices in settings of massive multiplicity has led to the development of the elimination by aspects model; etc. These models have produced immense

improvements in our understanding of choice data and in our ability to extract useful information about individual and aggregated preferences.

In contrast, the tradition of studying processes in psychology has been to focus on a bottom-up approach. Experiments in relatively simple settings have been carefully designed and conducted with the goal of elucidating features of the processes that underlie decision making. Plausible models for the processes have been proposed and compared to the experimental evidence. The match between the models and the experimental evidence is striking, suggesting that, in these settings, the posited processes provide a good description of the decision making process.

The research agenda that we have focused on is to take the basic processes and to extend them from relatively simple choice tasks to the complex choice tasks faced in marketing. The potential gains from such extensions are twofold: Better analysis of data sets, especially when it comes to transferring results from data collected in one situation to another context, and a better understanding of choice processes. This leads to several directions for research, some of which we have already begun to explore (Table 1).

First, experimental work aimed at clarifying basic features of typical choices encountered in marketing, such as serial versus parallel processing, the effect of workload on performance, the decisional stopping rules used, etc. will provide much needed guidance for modeling. The kinds of experiments we envision are based on methodologies such as the double factorial paradigm (e.g. Townsend & Nozawa 1995). The envisioned experiments are not aimed at simply identifying significant departures from standard models but translate into general structural requirements for better models in a particular but not singular context.

Choice problems in Marketing	
Experimental work	Modeling

Model free evidence
   - serial / parallel processing
   - processing capacity
   - decisional stopping rules
   - dependence structures
	Sequential sampling models
	   - model choices and other observables such as response time jointly
	   - model dependence among accumulation processes implied by particular processing styles (e.g. alternative based versus attribute based)
	   - explore connections to established choice models through reduced form representations / approximations
	   - formally extend the micro economic constrained maximization framework to include processing aspects
	   - develop practical algorithms and user friendly software

Table 1 Research issues

Second, sequential sampling models imply both choice probabilities and a distribution on time until a decision is made. This response time distribution and the choice are dependent, and this dependence implies that there is information in response time about an individual's preferences. This information should be collected as part of the choice experiment (and often is as a by product), and it should be used in the analysis.

Third, sequential sampling models can be extended to allow for dependence in the information accumulated toward particular choices. The creation of models that capture this dependence implied by an experimental, qualitative understanding of the choice process (e.g. serial processing of attributes as in Diederich 1997) in a structured fashion will make the models more realistic and should result in better choice analysis.

Fourth, the relationship between current top-down models and sequential sampling models needs to be better described. Understanding this relationship will suggest better top-

17

down approximations to sequential sampling models which may be easier to implement for practical analysis.

Fifth, a recurring theme of behavioral work is the appearance in nature of decision-making strategies which are optimal, or nearly optimal, subject to a set of constraints. A very recent stream of research investigates the optimality of some of the sequential sampling models in a constrained maximization framework (Bogacz et al. 2006, McMillen and Holmes 2005). Constrained maximization here refers to the question of what decision rules will lead to decisions with a given probability of error in minimal time or, nearly equivalently, what decision rule minimizes the amount of error given a fixed time. This work establishes (some) sequential sampling models not only as rich, quantitative descriptions of choice but as amenable to economic analysis.

Sixth, there is a need for software which can fit the process-based models or close approximations to them. The software would need to be efficient so that the models can be fit quickly, it needs to be able to handle data sets of realistic size and complexity, and it needs to be robust to handle vagaries of the data. This sort of software would facilitate not only analysis, but also design of the experiments used to collect the data.

References

Adamowicz, V., Bunch, D., Cameron, T. A., Dellaert, B. G. C., Hanneman, M., Keane, M., Louviere, J., Meyer, R., Steenburgh, T., & Swait, J. (2008). Behavioral Frontiers in Choice Modeling, *Marketing Letters*, this volume.

Ashby, F. G. (1983). A biased random-walk model for 2 choice reaction-times. *Journal of Mathematical Psychology, 27*, 277-297.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, 93, 154-179.

Bogacz, R., Brown, E., Jeff, M., Holmes, J.P., & Cohen, J.D. (2006). The Physics of Optimal Decision Making: A Formal Analysis of Models of Performance in Two-Alternative Forced-Choice Tasks, *Psychological Review*, 113, 700-65.

Bogacz, R. (2007). Optimal decision-making theories: linking neurobiology with behaviour. *Trends in Cognitive Sciences,* 11, 118-125.

Busemeyer, J.R. & Diederich, A. (2002). Survey of decision field theory. *Mathematical Social Sciences*, 43, 345 - 370.

Busemeyer, J. R. & Townsend, J.T. (1993). Decision field theory: A dynamic cognition approach to decision making. *Psychological Review*, 100, 432-59.

Diederich, A. (1997). Dynamic Stochastic Model for Decision Making under Time Constraints. *Journal of Mathematical Psychology*, 41, 260-74.

Diederich, A. (2003a). MDFT account of decision making under time pressure. *Psychonomic Bulletin and Review*, 10, 157 - 166 .

Diederich, A. (2003b). Decision making under conflict: Decision time as a measure of conflict strength. *Psychonomic Bulletin and Review*, 10, 167 - 176.

Diederich, A. & Busemeyer, J. R. (1999). Conflict and the stochastic dominance principle of decision making. *Psychological Science*, 10, 353-359.

Diederich, A., & Busemeyer, J.R. (2003). Simple Matrix Methods for Analyzing Diffusion Models of Choice Probability, Choice Response Time and Simple Response Time. *Journal of Mathematical Psychology*, 47, 304 - 322.

Gilbride, T. & Allenby, G. (2006). Estimating Heterogeneous EBA and Economic Screening Rule Choice Models. *Marketing Science.*

Huang, Y. & Hutchinson, J.W. (2008). Counting Every Thought: Implicit Measures of Cognitive Responses to Advertising. *Journal of Consumer Research*, 35, (June).

Huber, J. & Puto, C. (1983). Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects. *Journal of Consumer Research*, 10, 31-44.

Huber, J., Payne J.W., & Puto, C. (1982). Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis. *Journal of Consumer Research*, 9, 90-98.

Johnson, J. G. & Busemeyer, J. R. (2005). A dynamic, computational model of preference reversal phenomena. *Psychological Review*, 112, 841-861.

Kivetz, R., Netzer, O., & Srinivasan, V. (2004). Alternative Models for Capturing the Compromise Effect. *Journal of Marketing Research,* 41, 237–57.

LaBerge, D. (1962). A recruitment theory of simple behavior. *Psychometrika, 27*, 375-396.

McMillen, T. & Holmes, P. (2005). The dynamics of choice among multiple alternatives. *Journal of Mathematical Psychology*, 50, 30-57.

Otter, T., Allenby, G., & van Zandt, T. (2007). An Integrated Model of Choice and Response Time. Forthcoming at *Journal of Marketing Research*

Rieskamp, J., Busemeyer, J.R., & Mellers, B.A. (2006). Extending the Bounds of Rationality: Evidence and Theories of Preferential Choice. *Journal of Economic Literature*, 44, 631-661.

Roe, R. M., Busemeyer, J.R., & Townsend, J.T. (2001). Multialternative Decision Field Theory: A Dynamic Connectionist Model of Decision Making. *Psychological Review*, 108, 370-92.

Ruan, S. (2007). *Poisson Race Models for Conjoint Choice Analysis: Theory and Applications.* Unpublished Ph.D. dissertation. Department of Statistics, The Ohio State University.

Ruan, S., MacEachern, S., Otter, T., & Dean, A. (2007). Dependent Poisson Race Models and Modeling Dependence in Conjoint Choice Experiments. Forthcoming at *Psychometrika.*

Simonson, I. (1989). Choice Based on Reasons: The Case of Attraction and Compromise Effects. *Journal of Consumer Research*, 16, 158-174.

Smith, P.L. (2000). Stochastic Dynamic Models of Response Time and Accuracy: A Foundational Primer, *Journal of Mathematical Psychology*, 44, 408-436.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika,* 25, 251-260.

Tanner; M. A., & Wong, W. H. (1987). The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association*, 82, 528-540.

Townsend, J. T. (1972). Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, 25, 168-199.

Townsend, James T. and F. Gregory Ashby (1983), *Stochastic Modeling of Elemenatary Psychological Processes,* Cambridge University Press, Cambridge.

Townsend, J. T. & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial and coactive theories. *Journal of Mathematical Psychology*, 39, 321-360.

Townsend, J. T., & Schweickert, R. (1989). Toward the trichotomy method: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology*, 33, 309-327.

Townsend, J. T. & Wenger, M.J. (2004). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review*, 111, 1003-1035.

Tversky, A. (1972a). Elimination by aspects: A theory of choice. *Psychological Review*, 79, 281-299.

Tversky, A. (1972b). Choice by elimination. *Journbal of Mathematical Psychology*, 9 (4), 341-367.

Tversky, A. & Simonson, I. (1993). Context Dependent Preferences. *Management Science*, 39, 1179-1189.

Usher, M. & McClelland, J.L. (2004). Loss Aversion and Inhibition in Dynamical Models of Multialternative Choice. *Psychological Review*, 111, 757-769.

Vickers, D. (1970). Evidence for an accumulator of psychophysical discrimination. *Ergonomics*, 13, 37-58.