
3. Psychological research and theories of preferential choice

*Jared M. Hotelling, Jerome R. Busemeyer and
Jörg Rieskamp*

Understanding human preferential choice behavior is challenging because humans change their preferences across time and contexts. This chapter summarizes the basic behavioral findings from research on human preferential choice and reviews the psychological theories that have been proposed to account for puzzling findings. We focus on a class of theories that formalize the decision process as one of accumulating evidence to criterion. The main theme that we attempt to convey to the reader is that a coherent view of an individual's underlying beliefs and values can only be recovered by carefully modeling the dynamic nature of the choice process through which these beliefs and values operate to produce observed behavior.

When examining people's choice behavior, it becomes apparent that it varies substantially. For instance, Hey (2001) conducted a study in which 53 people repeatedly choose between pairs of simple gambles in five different sessions. In every session the same set of one hundred pairs was presented. When assuming stable and deterministic preferences all people should have made identical choices in every session. However, it turned out that no single person always made the same choices across all five sessions. Instead, on average participants changed their preferences for 10 percent of pairs between two consecutive sessions; for on average 23 percent of the pairs they did not make identical choices in all five sessions. The seminal work by Mosteller and Nogee (1951) discovered early on that people's choice behavior varies and has a probabilistic character. More surprisingly, its consequences for theory building are still not fully acknowledged. Axiomatic approaches to human choice behavior imply deterministic behavior and do not contain an error theory that could explain people's inconsistencies. The undeniable variability in people's behavior is often implicitly acknowledged by assuming that people's inconsistencies can be explained by unsystematic errors or "white noise." This implies that people's behavior varies around the theories' deterministic predictions. However, when people's inconsistencies are systematic such forms of "tremble error" theories (cf. Loomes et al., 2002) are not sufficient to explain human behavior. Moreover, without an explicit error theory it appears almost impossible to separate unsystematic from systematic inconsistencies and to unravel the mechanisms that underlie the systematic inconsistencies.

In this chapter we will describe the two standard approaches to explaining the probabilistic character of choice behavior represented by fixed and random utility models. Second, we will summarize the key empirical findings violating essential principles of utility theories. Next, we will present dynamic psychological models of choice behavior that describe how people's preferences evolve over time, and that have been suggested to explain some of the observed behavioral (ir)regularities. Then we briefly compare the

different approaches and models against each other, and finally we conclude by summarizing some additional and new directions.

1 CLASSIC PROBABILISTIC UTILITY MODELS

Before delving into the behavioral research on preferential choice, it is useful to first spell out two classic utility theory approaches: *fixed* versus *random* utility models. These theories served as a primary guide for past research on preferential choice by implying specific choice principles that were tested empirically. These empirical tests revealed the inadequacies of the theories in describing people's behavior and led to the development of more descriptive and cognitively driven models. In the following we will describe and define these models using the notation and terms employed earlier in Rieskamp et al. (2006).

Assume that there exists a complete set of choice options $X = \{A_1, \dots, A_n\}$ under consideration. The person may be presented with a subset $Y = \{B_1, \dots, B_m\} \subseteq X$, $m \leq n$ of this complete set. The probability that an individual chooses option B_i from the set Y is denoted $p(B_i|Y)$, with the constraint $p(B_i|Y) \geq 0$ and $\sum_{i \in Y} p(B_i|Y) = 1.0$. Note that psychological models normally assume that an individual's behavior is probabilistic, and so the theory needs to be defined with the probabilities at an individual (pooled across replications within an individual) rather than an aggregate level (pooled across individuals).

1.1 Fixed Utility Approach

According to the *fixed utility model*, a real value $u(A_i)$ can be assigned to each option $A_i \in X$ that remains fixed across choice sets Y . When presented with the subset Y , the probability of choosing option B_i equals

$$p(B_i|Y) = f(u(B_i), u(B_1), \dots, u(B_{i-1}), u(B_{i+1}), \dots, u(B_m)), \quad (3.1)$$

where the function f is a strictly increasing function of the first coordinate, $u(B_i)$, and a strictly decreasing function of each of the remaining coordinates. For example, according to a Luce (1959) ratio of strength model

$$p(B_i|Y) = \exp(u(B_i)) / \sum_{j \in Y} \exp(u(B_j)). \quad (3.2)$$

A key idea of this class of models is that choice is inherently probabilistic and fundamentally unpredictable. Even with fixed utilities, a person's choice on each occasion remains probabilistic. The main simplifying assumption of this model is that the utilities assigned to the choice options do not depend on the choice set Y . In other words, utilities are context independent, also called simple scalable (Luce and Suppes, 1965). This context independence property produces a shortcoming of this class of models, namely that effects of the context on the choice situation can hardly be explained by fixed utility models. Naturally, this property can be relaxed, but how to do this in a coherent and parsimonious manner is quite challenging.

1.2 Random Utility Approach

According to *random utility models*, on any choice occasion, the person samples an evaluation point ω from a sample space Ω that determines the n real valued utilities $U(\omega) = [U_1(\omega), \dots, U_n(\omega)]$. This sampling process produces n random variables U_i , $i = 1, \dots, n$ that determine an n dimensional random utility distribution function denoted by $F_X = \Pr[U_1 \leq u_1, \dots, U_n \leq u_n]$; when presented with a choice set Y , the person chooses the option that has the maximum randomly sampled value, $\max\{U(B_1), \dots, U(B_m)\}$. The choice probabilities for the set Y are based on the marginal m dimensional distribution F_Y , which is obtained by integrating F_X over the values of the options in X that are not presented in Y . The probability that option B_i is chosen from the presented set Y equals the probability that the randomly sampled value for the random variable U_i is the maximum:

$$p(B_i|Y) = \Pr[U_i = \max\{U_j, U_j \in Y\}]. \quad (3.3)$$

For example, the Thurstone (1959) model assumes that the distribution function for the random utility vector $U = [U_1, \dots, U_n]$ is multivariate normal.¹ The key idea behind this class of models is that once the evaluation point ω is selected by the person, then choice becomes deterministic. Behavior is only probabilistic because we do not know the point ω used to evaluate the options; we only know the probabilities of sampling these points. The main assumption of this class of models is that the same n dimensional distribution function F_X is used for all choice sets Y . In other words, the distribution function is context independent. This context independence turns out to produce shortcomings for explaining certain choice phenomena with this class of models. As with fixed utility models, this property can be relaxed, but doing so in a coherent and parsimonious manner remains a major challenge.

1.3 Comparison of the Two Utility Approaches

The ideas motivating these two approaches are fundamentally different: Fixed utility models imply that choice is fundamentally indeterministic. In contrast, random utility models imply that choice is completely deterministic. Despite these conceptual differences, it is often difficult to empirically discriminate these ideas. For example, the Luce (1959) model can be mathematically derived from a random utility model with an identical and independently extreme value distributed error (Yellot, 1977). More general extreme value random utility models form the basis of many economic choice models (McFadden, 1981), and they are also very popular among marketing researchers (Louviere et al., 2000). In sum, fixed and random utility models represent the predominant approach (in particular, in economics) of predicting human choice behavior. Both approaches should be evaluated empirically by the major behavior findings.

2 BASIC CHOICE BEHAVIOR FINDINGS

This section reviews the basic empirical findings from preferential choice that have accumulated across the last 50 or so years from behavioral economics, consumer research, and psychology. As mentioned earlier, much of this research was targeted at basic properties implied by the two classic probabilistic utility approaches described above. The research examines whether human choice behavior obeys basic properties such as transitivity, independence, regularity, and stationarity.

2.1 Transitivity

One of the most basic properties of choice to examine empirically is transitivity. If a person prefers Beethoven to Mozart and Mozart to Chopin, then transitivity implies that the person will also prefer Beethoven to Chopin. It appears unreasonable to violate this principle repeatedly.

In general, the transitivity property is considered to be one of the main axioms of rational choice. Formally, transitivity is defined by a mathematical relation called a “preference” denoted \succsim so that $B_i \succsim B_j$ means that option B_i is preferred or indifferent to option B_j . Transitive preferences must satisfy $B_i \succsim B_j \wedge B_j \succsim B_k \rightarrow B_i \succsim B_k$ for all i, j, k in X . This is required by deterministic utility models in order to postulate a real valued utility function for ordering preference, $u: X \rightarrow \text{Reals}$.

Transitivity is often justified as an axiom of choice by arguing that violations of this axiom permit a person to be turned into a money pump. If a person prefers B to C , then the person should be willing to pay money to exchange C for B ; likewise, if the person prefers A to B then the person should be willing to pay money to exchange B for A ; finally if the person also prefers C to A then the person should be willing to pay to exchange A for C , thus returning to their original position, but after losing money on three exchanges. However, it is unlikely that these intransitive preferences could be exploited by building a money pump (cf. Chu and Chu, 1990). Instead, people will presumably notice their intransitive cycles at some point, making the money pump a “bogeyman” that only demonstrates *in principle* the irrationality of intransitive choices but one that would never be observed (Lopes, 1996).

The concept of transitivity is difficult to apply to probabilistic choice behavior because it is difficult to define the preference relation \succsim when choices are inconsistent. One way is to define $B_i \succsim B_j \leftrightarrow p(B_i | \{B_i, B_j\}) \geq .50$ (Luce, 2000). This immediately leads to a definition of *weak stochastic transitivity*:

$$p(B_i | \{B_i, B_j\}) \geq .50, p(B_j | \{B_j, B_k\}) \geq .50 \rightarrow p(B_i | \{B_i, B_k\}) \geq .50 \quad (3.4)$$

for all i, j, k in X .

The fixed utility class of models must satisfy weak stochastic transitivity, and moreover, this class must also satisfy a stronger version called *strong stochastic transitivity*:

$$p(B_i | \{B_i, B_j\}) \geq .50, p(B_j | \{B_j, B_k\}) \geq .50 \rightarrow p(B_i | \{B_i, B_k\}) = \max \{ p(B_i | \{B_i, B_j\}), p(B_j | \{B_j, B_k\}) \} \text{ for all } i, j, k \text{ in } X. \quad (3.5)$$

This follows from the fact that $p(B_i|\{B_i, B_j\}) \geq .50 \rightarrow u(B_i) \geq u(B_j)$ and $p(B_j|\{B_j, B_k\}) \geq .50 \rightarrow u(B_j) \geq u(B_k)$ and together this implies $u(B_i) \geq u(B_k)$, so that $p(B_i|\{B_i, B_k\}) \geq p(B_j|\{B_j, B_k\})$ and $p(B_i|\{B_i, B_k\}) \geq p(B_i|\{B_i, B_j\})$.

The random utility class of models does not need to satisfy weak stochastic transitivity. The decision maker can be transitive from each point of view ω , but averaging across these different preference orders produced by different points of view can violate weak stochastic transitivity (Regenwetter et al., 2011). For instance, suppose that across different occasions, a person experiences the following three transitive preference orders: $(B_1 > B_2 > B_3)$, $(B_2 > B_3 > B_1)$, $(B_3 > B_1 > B_2)$ where $A > B$ indicates strict preference. If we assume that these preferences occur equally often, then we observe that $p(B_1|\{B_1, B_2\}) = 2/3$, $p(B_2|\{B_2, B_3\}) = 2/3$, but $p(B_1|\{B_1, B_3\}) = 1/3$, violating weak stochastic transitivity, which is called the *Condorcet paradox*.

Although random utility models do not need to satisfy weak stochastic transitivity, they need to satisfy another transitivity property called the triangular inequality (although unrelated to the distance axiom) for binary choices (Marschak, 1974), if we assume no indifference, $p(U_i = U_j) = 0$:

$$p(B_i|\{B_i, B_j\}) + p(B_j|\{B_j, B_k\}) - p(B_i|\{B_i, B_k\}) \leq 1. \quad (3.6)$$

This can be shown as follows. For brevity, define $p(xyz)$ as the probability for the strict transitive order $x > y > z$. Then $p(x|\{x, y\}) + p(y|\{y, z\}) - p(x|\{x, z\}) \leq 1 \rightarrow (1 - p(y|\{x, y\})) + (1 - p(z|\{y, z\})) - (1 - p(z|\{x, z\})) \leq 1 \rightarrow p(z|\{x, z\}) \leq p(z|\{y, z\}) + p(y|\{x, y\}) \rightarrow p(zxy) + p(zyx) + p(yzx) \leq p(zyx) + p(zxy) + p(xzy) + p(yxz) + p(yzx) + p(zyx)$. (Note that $p(z|\{x, z\}) = p(zxy) + p(zyx) + p(yzx)$, because the right hand side consists of the three rank orders that produce a choice of z over x).

Given the importance of the transitivity property, one would expect a firm resolution with regard to its empirical status. Beginning with May (1954), a long series of investigations has appeared reporting violations of transitivity (see review by Rieskamp et al., 2006). Many of these studies were designed to replicate the well-known experiment by Tversky (1969), who reported violations of weak stochastic transitivity. Recently, however, the results regarding transitivity have been called into question because of inadequate methods for statistically testing this property (Iverson and Falmagne, 1985; Regenwetter et al., 2011). In response, several new statistical methods have been developed for testing transitivity, and these send a more mixed message. Tsai and Böckenholt (2006) used a Thurstone random utility mixture model to compare a parametric formulation of an intransitive model versus a more constrained transitive model, and they found that chi-square difference tests rejected the constrained transitive model in favor of the more general intransitive model. Regenwetter et al. (2011) developed nonparametric statistical tests of the triangular inequality and found that most participants did not produce statistically significant violations. Birnbaum and Schmidt (2008) developed statistical models for choice patterns based on true preferences plus response errors, and they found that most participants produced patterns consistent with a model of transitive true preferences plus response error. In sum, in spite of the large body of research it appears unclear whether people violate weak stochastic transitivity. Recent evidence using adequate statistical tests shows that these violations appear to be the exception rather than the rule.

More systematic, reliable, and robust evidence has been found for violations of *strong stochastic transitivity* (Mellers and Biagini, 1994; Rumelhart and Greeno, 1971; Tversky and Russo, 1969). The psychological reason for the violations of strong stochastic transitivity is the following: when a person is faced with choices between multidimensional options, the *context* produced by a *pair* of options is used to single out some dimensions for making comparisons while ignoring other dimensions. For example, one pair of options A and B might differ largely in a tradeoff among the quality features but seem similar in price – for this pair the person may tend to ignore the price difference and focus more on the quality features. In another example comparing B and C again the price difference appears negligible after giving the quality dimension more attention. However, when comparing A and C the difference in price might become substantial, so that the person may tend to refocus more heavily on price for the comparison. Thus, the dimensions used to make the comparison change across choice pairs and this change in the basis for comparison results in violations of strong stochastic transitivity.

2.2 Independence

The next most basic property of choice to be examined empirically is independence from irrelevant alternatives, which is concerned with invariance of a preference relation between two options with changes in the choice set that contains these options. More formally suppose $Y \subseteq X$ and also $Z \subseteq X$; then independence states (Tversky, 1972b)

$$p(B_i | \{B_i, B_j\} \cup Y) \geq p(B_j | \{B_i, B_j\} \cup Y) \leftrightarrow p(B_i | \{B_i, B_j\} \cup Z) \geq p(B_j | \{B_i, B_j\} \cup Z). \quad (3.7)$$

This property is also required by fixed utility theories: $p(B_i | \{B_i, B_j\} \cup Y) \geq p(B_j | \{B_i, B_j\} \cup Y) \rightarrow u(B_i) \geq u(B_j) \rightarrow p(B_i | \{B_i, B_j\} \cup Z) \geq p(B_j | \{B_i, B_j\} \cup Z)$. However, this property is not required by the general random utility model (although the standard multinomial logit random utility model with statistically independent utilities satisfies this property). The independence property directly reflects the context independence property underlying the fixed utility model. There are different ways in which this principle can be systematically violated in empirical studies. One very prominent way is the so-called *similarity effect*, and another is the *compromise effect*.

The similarity effect was initially examined by Tversky (1972a). The essential idea is to examine the preferences for options, say A versus B , with or without the context of another option A^* . Options A and B are designed to be qualitatively different or dissimilar – for example option A could be high in quality and price whereas option B is low in quality and price. Option A^* is designed to be very similar and competitive with option A , for example it could also be high in quality and price but just a bit worse in quality but therefore a bit cheaper. It has been empirically found that in binary choice it can be that A is preferred to B , $p(A | \{A, B\}) \geq p(B | \{A, B\})$, but when option A^* is included into the choice set then A^* competes with A but not so much with B , so that $p(A | \{A, B, A^*\}) \leq p(B | \{A, B, A^*\})$. These findings were later extended by Tversky and Sattath (1979). In the transportation literature the similarity effect is often explained by the red-bus blue-bus example, where the decision maker chooses between different transportation options, such as a car and a red bus. When another (blue) bus is added to the choice set that only differs in color with the red bus it harms the choice share of the red bus but not of the car (Train, 2003).

The compromise effect was initially examined by Simonson (1989). Again, the essential idea is to examine the preferences for options, say B versus C with or without the context of another option A . In this case, option C is designed to be a compromise between the two other options A and B , and the latter are designed to have extreme tradeoffs on two different dimensions. For example, option A may be very high in quality and price whereas option B is very low in quality and price, and option C is moderate in quality and price. In the context of a choice between B and C , option C does not appear as any kind of compromise – option B is simply lower than C in quality and price; but if another extreme option A is included in the choice set, then option C emerges as a compromise that lies between the two extremes of A and B . The main finding is that when comparing options B and C one finds $p(B|\{B,C\}) \geq .50$; but when option A is included into the choice set then the compromise becomes favored so that $p(C|\{A,B,C\}) \geq p(B|\{A,B,C\})$. This work has been replicated and extended in many subsequent studies (see review by Kivetz et al., 2004).

2.3 Regularity

Another very basic property of choice to be examined empirically is regularity, which asserts that the addition of an option to a choice set should never increase the probability of selecting an option from the original set. More formally, suppose $Z \subset Y \subseteq X$ presented for choice. Then regularity states

$$p(B_i|B_i \cup Z) \geq p(B_i|B_i \cup Y). \quad (3.8)$$

This property holds for random utility theories since it is less likely that B_i has the highest utility in a larger set, Y , than in a smaller set, Z , because $p(B_i|B_i \cup Z) = \Pr[U(B_i) = \max\{U(B_j), B_j \in Z\}] \geq \Pr[U(B_i) = \max\{U(B_j), B_j \in Z\}] \times \Pr[U(B_i) = \max\{U(B_k), B_k \in Y - Z\} | U(B_i) = \max\{U(B_j), B_j \in Z\}] = p(B_i|B_i \cup Y)$.

Violations of regularity were first found by Huber et al. (1982), which are called *asymmetric dominance effects* and/or *attraction effects*. The essential idea here is to examine the preferences for options, say A versus B with or without the context of another *deficient* option D . Like the similarity effect, options A and B are designed to be qualitatively different or dissimilar – for example option A could be high on quality and price whereas option B is low in quality and price. Also like the similarity effect, option D is designed to be very similar to option A ; but the *key* difference needed to produce an attraction effect is to make D deficient or defective as compared to option A , for example D is of slightly lower quality but the same price as A so that D is dominated by A . In this situation, D is rarely chosen, but including D in the choice set increases the probability of choosing option A so that $p(A|\{A,B\}) \leq p(A|\{A,B,D\})$. The same result can be obtained even when D is not necessarily dominated by A but is just much less attractive as compared to A (Huber and Puto, 1983). These findings have been replicated in many studies (Wedell, 1991). For review see Heath and Chatterjee (1995).

2.4 Stationarity

All choices take time, and the time taken to make a decision can change the choice that is finally made. For example, suppose the reader wakes one morning to find two emails, each of which is an invitation to present a keynote speech at an attractive venue, but unfortunately on the same day. Choosing between these mutually exclusive offers may take substantial time to think through the advantages and disadvantages, and deadlines for making the decision could affect the final decision by preventing one from thinking through all the consequences.

If choice probabilities do not change as a function of deliberation time (excluding time to read the choices), then they are stationary. Define the probability of choosing an option A from a set $Y \subset X$ conditioned on deliberating for a period of time t as $p_t(A|Y)$. Stationarity states that $p_t(A|Y) = p(A|Y)$ for $t > t_0$ where t_0 is time necessary to read the choice options. Fixed and random utility models are static models that provide no mechanisms for predicting the effects of decision time on choice probability. This oversimplification becomes a problem for these theories when stationarity is violated. On the one hand, one could argue that fixed and random utility models only apply for choices without time constraints, assuming that the probabilities converge to some asymptote $p_t(A|Y) \rightarrow p_\infty(A|Y)$. On the other hand, even when there is no explicit deadline, there is a cost for taking time to deliberate that puts time pressure on the decision maker.

The effects of decision time on choice probability are now well established by several different lines of experimental research (Svenson and Maule, 1993). Decision makers often become more inconsistent under time pressure (Olschewski and Rieskamp, 2021). Consumer choices systematically reverse under time pressure (Svenson and Edland, 1987). Choices between uncertain actions systematically change as a function of deliberation time (Busemeyer, 1985) and even reverse under time pressure (Diederich, 2003a, 2003b). Compromise and attraction effects become even larger when decision makers are encouraged to deliberate longer (Dhar et al., 2000). Decision makers also tend to switch strategies for making decisions when pressed for time (Ben Zur and Bresnitz, 1981; Rieskamp and Hoffrage, 2008; Olschewski and Rieskamp, 2021).

Time is an important factor in choice behavior because information about the choice options must be accumulated across time, and the type of information ultimately entering the choice process depends on how much time is allocated to making the decision (Wallsten and Barton, 1992). Under short time constraints, the decision maker may have time only to focus on the most important dimension and ignore a number of other relevant aspects of the decision problem (e.g., Rieskamp and Hoffrage, 2008). When given more time to process all of the dimensions, the initial preference established by the first dimension can be overcome by competing information accumulated on many other relevant dimensions. Alternatively, decision strategies could change from optimal to heuristic under time pressure (Payne et al., 1993).

Clearly, decision time plays a major role in emergency types of decisions, such as medical or military decisions (Janis and Mann, 1977). However, decision time also plays an important role in day-to-day economic choices by consumers because of the well-known tradeoff between effort and accuracy when choosing strategies for decision making (Payne et al., 1993). Time has a cost and consumers are often unwilling to spend a large amount of costly time to make decisions (Wright, 1972).

There is another way to think about stationarity, one that concerns changes in the choice probabilities caused by learning from experience across a sequence of repeated decisions. There is now a large experimental literature on experienced-based choice (cf. Hertwig and Erev, 2009), and a growing theoretical effort to develop integrated models of learning and decision making (Erev and Barron, 2005; Gonzalez and Dutt, 2011; Fontanesi et al., 2019). However, the topic of learning goes beyond the intended range of this chapter.

2.5 Conclusions from Behavioral Findings

The overview of some of the main findings on human choice behavior illustrates that people often do not adhere to simple principles such as strong stochastic transitivity, independence, regularity, and stationarity. People's preferences and choice behavior change as a function of the choice context as well as the amount of time taken to process the information for making a decision. However, standard fixed and random utility models are insufficiently sensitive to context or time pressure to capture these systematic changes in preference behavior. For instance, fixed utility theories assign values to options that are independent of the choice set, and random utility theory do not have any mechanism to explain the impact of available time on decision making. Of course, one can change utilities in an ad hoc way for every context to fit these effects post hoc. For instance, the utility function could be defined to include a context effect term (e.g., include an extra utility term for a compromise option), and different random utilities could be assumed for different deliberation times (e.g., change the means for each time period). But these ad hoc fixes do not provide a scientific explanation that allows for a priori predictions in new contexts and time constraints. Psychological choice models, which are described next, tend to be more complex than the basic fixed and random utility models. However, they are designed to provide mechanisms to account for either context effects, or time pressure effects, or both.

3 DYNAMIC PSYCHOLOGICAL MODELS

Traditional psychological models do not explicitly describe how people's preferences change as a function of deliberation time when there is a specific deadline. Nor do they describe the random time, T , chosen by the decision maker to make a decision when there is no specific deadline. Here we review several prominent dynamic models of choice behavior. A central feature of these is that they explicitly model how a deadline decision time affects people's choices, as well as the time T required to make a decision when there is no specific deadline. By focusing on the psychological processes underlying decisions, these models provide a plausible account of choice behavior across a wide variety of situations and domains. As a result, theories of evidence accumulation have been used to account for myriad phenomena, such as perceptual decisions (e.g., Dutilh and Rieskamp, 2016; Ratcliff and Rouder, 1998, 2000), inference (Lee and Cummins, 2004; Trueblood et al., 2014), pricing (Johnson and Busemeyer, 2005), intertemporal choice (Gluth et al., 2017), consumer choice (Krajbich et al., 2012), planning (Hotaling, 2020), memory (Cox and Shiffrin, 2017), and classic decision biases (Hotaling et al., 2015).

3.1 Horse Race Choice Models

One of the first attempts to generalize the classic random utility family to include choice response time was a class called *horse race* random utility models (Marley and Colonius, 1991). The underlying assumption is that when presented with a choice from set Y , each option is assigned a non-negative arrival time and option $B_i \subseteq Y$ is chosen if its “arrival event” occurs first. It is assumed that the random arrival times for all options in the entire set X form a joint multivariate distribution, and the probability distribution for a subset $Y \subseteq X$ is obtained by marginalizing over the joint multivariate distribution of arrival times for the complete set X . The probability of choosing option B_i from set Y and this occurs after time t equals

$$p(B_i \wedge (T > t) | Y) = p([T(B_i) = \min\{T(B_j), B_j \in Y\}] \cap T(B_i) > t). \quad (3.9)$$

The horse race choice model reduces to a random utility model when marginalizing over the time to make the decision; however, it extends the traditional random utility model by also providing a model of the distribution of choice time T .

Building on earlier ideas presented in Townsend and Ashby (1983), Otter et al. (2008) formulated an “independent” version of the horse race model. Otter et al. (2008) assumed that each choice option B_i in a set Y is associated with an independent counter $N_i(t)$ that counts up events favoring that option at time t . The first counter to reach a threshold wins the race and determines the choice. Events favoring an option occur at times distributed according to a Poisson process, and the rate assigned to each option depends on the attribute values of the choice option. The model was fit to choice and response time data from a survey involving 422 people, with each person providing stated choices for 18 different choice sets, and each set contained five TV sets described by six attributes. The researchers found that by fitting both choice and response time (as compared to fitting only choice), they obtained a better marginal fit to the choice part of the data, which suggests that the model successfully extracts information contained in response times and that response times are informative about the cognitive processes underlying the observed choices.

The horse race choice model is a random utility model, and so it inherits the triangle inequality and regularity properties. Therefore, it cannot explain violations of regularity such as the attraction effect. It has not been used to account for similarity effects, nor is it clear how it could ever explain the compromise effect. The independent Poisson horse race model does account for changes in choice probability as a function of deliberation time. However, the latter model also predicts that the distribution of choice times becomes normal as the counter threshold increases (Otter et al., 2008). This prediction is problematic because, in fact, the choice response time distribution tends to become more positively skewed with longer mean deliberation times (Ratcliff and Smith, 2004). This problem is corrected by using the next class of dynamic models.

3.2 Sequential Sampling Choice Models

Sequential sampling models of decision making were originally developed for Bayesian inference (DeGroot, 1970). Cognitive psychologists applied these models to a variety

of cognitive tasks including sensory detection (Smith, 1995), perceptual discrimination (Laming, 1968; Link and Health, 1975; Usher and McClelland, 2001; Vickers, 1979), memory recognition (Brown and Heathcote, 2005; Ratcliff, 1978); categorization (Ashby, 2000; Nosofsky and Palmeri, 1997), and probabilistic inference (Wallsten and Barton, 1982). Several sequential sampling models for preferential choice have also been proposed (Aschenbrenner et al., 1984; Bhatia, 2013; Fehr and Rangel, 2011; Glöckner and Betsch, 2008; Guo and Holyoak, 2002; Roe et al., 2001; Usher and McClelland, 2004).

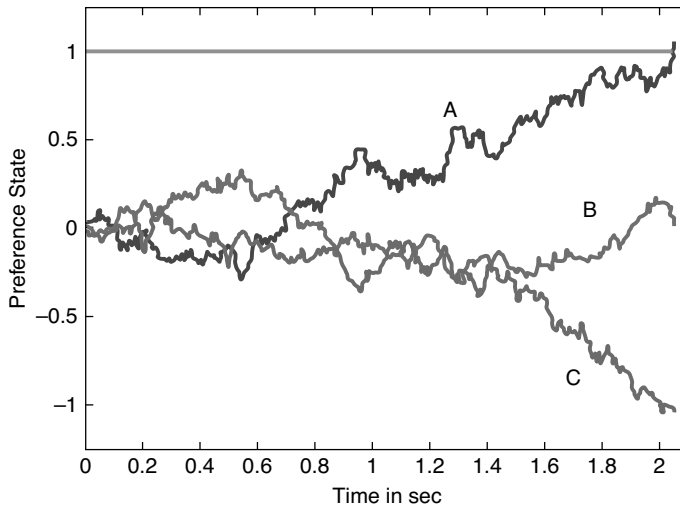
The basic idea of many sequential sampling models is that when presented with a subset of m choice options $Y \subseteq X$ from a larger set X , a choice from this set takes some amount of deliberation time T . The decision process starts with an $m \times 1$ state vector $P(0)$, where each coordinate corresponds to the preference for one of the options in Y . This initial state $P(0)$ reflects preferences for status quo options or biases from past experience. The initial state vector then evolves across time t by accumulating evaluations regarding the advantages and disadvantages of each option. After deliberating for a time t , the cumulative evaluations evolve to a new state vector $P(t)$ according to the linear stochastic difference equation (technically an Ornstein-Uhlenbeck (OU) process),

$$dP(t+h) = P(t+h) - P(t) = -\Gamma \cdot P(t) \cdot h + \mu \cdot h + dB(t+h), \quad (3.10)$$

where $dB(t+h) = B(t+h) - B(t)$ is a Brownian motion increment with mean zero and variance-covariance matrix $\Phi \cdot h$. The term, $\mu \cdot h + dB(t+h)$, is the new sample input into the process during the time step h , and $-\Gamma \cdot P(t) \cdot h$ is a feedback process that can be used to maintain preference stability. An important special case is the Wiener process, which is obtained by setting $\Gamma = \mathbf{0}$, but this allows preferences to grow without bound, and it also reduces the capability of accounting for context effects. The preference accumulation process continues until one of two stopping rules is satisfied – either a *fixed* or a *variable* stopping rule.

According to a *fixed* time stopping rule, the preference state evolves until some externally determined deadline time t occurs. In this case, the process evolves in an unconstrained manner until the deadline, t , at which point the option with the strongest preference state is chosen (i.e., the option corresponding to the maximum coordinate of the $P(T)$ vector). In the case of the fixed time stopping rule, the theory is equivalent to a dynamic version of a Thurstone model in which the state vector $P(t)$ serves as an evolving vector of random utilities. However, unlike the static Thurstone model, the mean and the variance-covariance matrix of $P(t)$ evolve across time, which accounts for changes in preferences as a function of deliberation time. Furthermore, if the feedback matrix is non-zero ($\Gamma \neq \mathbf{0}$), then both the mean and variance-covariance matrix of $P(t)$ are context-dependent; that is, the multivariate distribution of $P(t)$ changes depending on the presented choice set Y . More specifically, if $\Gamma \neq \mathbf{0}$, then the joint distribution for a set of options Y cannot be derived by marginalization from a complete joint distribution defined over all possible options in X . Consequently, these theories do not have to satisfy regularity.² If the initial state $P(0)$ starts out unbiased (i.e., $P(0) = \mathbf{0}$) then these theories satisfy both weak stochastic transitivity and the triangle inequality. However, they generally violate strong stochastic transitivity because of the changes in the variance-covariance matrix Φ across pairs of options.

According to the *variable* time stopping rule, preferences continue to evolve until the preference strength for one of the coordinates, corresponding to one of the options, exceeds a positive threshold. The first option exceeding the threshold is chosen, and the deliberation time equals the time it takes for an option to first cross a threshold. Figure 3.1 illustrates the process using the variable stopping time for three options {A, B, C}. In this figure, the horizontal axis represents time, and the vertical axis represents the preference state for each option; and in this case option A first crosses the threshold at a time equal to 2 seconds. The flat line near the top of the figure is called the threshold, and it is used to determine how strong the preference must be in order to make a decision. Increasing this threshold increases the average time to make a decision, which allows more evaluations of all of the options' advantages and disadvantages. Decreasing the threshold decreases the average time to make a decision, which limits the ability to evaluate all the advantages and disadvantages of each option. The threshold is determined by the cost of sampling new information as compared to the gains and losses expected by the final decision after sampling more information. When the variable stopping time rule is used, sequential sampling models can be used to predict both the choice probabilities as well as the mean decision times. In fact, a strong test of the model is obtained by fitting the model parameters to the choice probabilities, and then using these same parameters to predict mean decision time. Overall, sequential sampling models provide very accurate accounts of the skewed shape of choice response time distributions (Ratcliff and Smith, 2004).



Note: The horizontal axis represents time, the vertical axis represents preference state for each option, and the flat horizontal line represents the decision threshold that has to be crossed by one of the preference states so that a choice occurs. In this case, option A first crosses the threshold at a time near to 2 seconds and is chosen by the decision maker.

Source: Adapted from Roe et al. (2001).

Figure 3.1 The preference state for each option/action changes over time by sampling information about the choice options

3.3 Decision Field Theory

As mentioned above, there are several versions of this general type of sequential sampling theory of preferential choice. One of the earliest was decision field theory (DFT, Busemeyer and Townsend, 1993; Roe et al., 2001). DFT shares exactly the same parameters as the traditional Thurstone model with respect to the mean vector μ and the covariance matrix Φ of the OU process. According to DFT, the mean μ and covariance matrix Φ are determined by the expectation and variance–covariance, respectively, of advantages or disadvantages for each option along a randomly sampled attribute (Roe et al., 2001). Different attributes are sampled across time by switching attention from one attribute to another during deliberation. Also key to DFT, the coefficients in its feedback matrix Γ are determined by the distance between options in a multi-attribute space of options (Hotelling et al., 2010). Three additional parameters are required to determine the feedback matrix Γ on the basis of the distances between the m options in the set Y (see Hotelling et al., 2010, for details).

3.4 Leaky Competing Accumulator

The leaky competing accumulator (LCA) model is another important version that shares many assumptions with DFT, but differs on some key mechanisms (Usher and McClelland, 2004).³ This model also assumes that people switch their attention during the decision process to the different attributes and accumulate advantages or disadvantages for each option across time. Note that DFT and LCA are competitors because they use different mechanisms to account for major behavior findings. Nonetheless, these models are quite similar and share many assumptions in contrast to more traditional random utility theories (Tsetsos et al., 2010).

3.5 Attention Drift Diffusion Model

Krajbich et al. (2010) developed a version of sequential sampling called the attention drift–diffusion model (ADD). Unlike DFT and LCA, it does not assume attention switching across attributes, and instead it assumes attention shifting across choice alternatives. The ADD assumes that each option accumulates random utilities according to a Wiener process ($\Gamma = 0$). The mean drift rate μ is simply determined by a traditional weighted sum of the attribute values, and $\Phi = \sigma^2 \cdot \mathbf{I}$, where σ^2 is the variance (a scalar). However, the mean drift rate is not constant during deliberation, but changes depending on the option to which the decision maker is attending. Attention is operationalized using eye-movement recordings, which track the location of the decision maker's gaze at each moment during the decision process. The mean drift rate of the Wiener process is assumed to change depending on the gaze of the decision maker, with the currently viewed option being given more weight. If the option under gaze is advantageous then its advantage is enhanced during the gaze. If the option under gaze is disadvantageous, then this disadvantage is enhanced during the gaze.

3.6 Associative Accumulator Model

The associative accumulation model (AAM, Bhatia, 2013), like DFT and LCA, assumes that decision makers attend to one attribute at a time, sequentially accumulating advantages and disadvantages for options based on attention switching across attributes. However, unlike DFT and LCA, the model assumes that the probability of attending to an attribute depends on the options in the choice set. Attributes with a stronger association to options receive greater activation and are more likely to be attended to.

3.7 Multi-Alternative Linear Ballistic Accumulator Model

The multiattribute linear ballistic accumulation model (MLBA, Trueblood et al., 2014) involves two components. A front-end first transforms attributes into subjective values, then evaluates alternatives relative to each other in a pair-wise fashion, and applies attention weights as an exponentially increasing function of similarity in the attribute space. MLBA assumes that comparisons between similar attributes receive greater attention, and that similarity may be asymmetric. These front-end processes produce context-dependent drift rates that serve as input to a back-end component comprised of the linear ballistic accumulator model (Brown and Heathcote, 2008), a type of deterministic horse race model.

3.8 Summary of Dynamic Models

The dynamic psychological models are more complex than the static psychological models, but they substantially expand their explanatory power. First, they are capable of predicting both choice probability and the distribution of time taken to make a choice. They are also capable of predicting the effects of time limits or time pressure on choice probabilities. Third, the dynamics provide mechanisms for simultaneously explaining all three types of context effects (similarity, attraction, compromise) as well as many other violations of independence and strong stochastic transitivity. None of the previously developed static models were found to be capable of predicting all these effects. This does not mean that it is not possible to discover a new static model that can do the job, but such a model remains to be developed.

Dynamic models, and sequential sampling models in particular, make strong predictions about the sampling process used to make decisions. One of the basic assumptions (except for the ADD) is that information sampling is not affected by the person's current preference (formally, the input, $\mu \cdot h + dB(t+h)$, does not depend on the current state of preference $P(t)$). However, recent studies using information search (Willemsen et al., 2011) and eye movements (Fiedler and Glöckner, 2012) indicate that the sampling process becomes biased to sample attributes that support the currently favored alternative. This finding suggests that the current preference state may feedback and modify the input into the accumulation process (see, e.g., Guo and Holyoak, 2002; Glöckner and Betsch, 2008).

4 COMPARING MODELS

Here we compare how well the models account for various behavioral phenomena. We focus our summary on key findings from the decision-making literature where several models have been applied.

4.1 Similarity Effect

Several models have been applied to the similarity effect. DFT, LCA, and AAM – which share the assumption that decision makers sequentially accumulate advantages and disadvantages for options based on comparisons within attributes – use attention switching across attributes to account for similarity effects and violations of independence and strong stochastic transitivity. For MLBA, asymmetric attention weighting produces similarity effects because greater weight is placed on positive evidence than negative evidence, and this asymmetry grows as options become more dissimilar. PR and ADD have not been used to study similarity effects, so it remains unclear if they can account for these behavioral patterns.

4.2 Attraction Effect

Competing explanations for the attraction effect come from several models. For DFT, the dominated option will tend to have negative activation, which feeds into and boosts the activation of the other options through lateral inhibitory connections. Since inhibition is distance dependent, the similar option benefits more than the dissimilar, resulting in the attraction effect. LCA takes a different approach. Rather than assuming distance-dependent inhibition between alternatives, it incorporates the same loss aversion principle used in the context dependent preference model (Tversky and Simonson, 1993) to predict attraction effects. AAM and MLBA likewise do not make use of the feedback matrix Γ ($\Gamma = 0$), but instead account for attraction effects in their front-end, before deliberation begins. For AAM, the average attention weight given to an attribute mainly depends on the total values assigned to the attribute across alternatives. Attraction effects arise because weights favor the stronger attribute for the dominating alternative. In MLBA, similarity effects arise because comparisons between similar options receive greater attention, which benefits the dominant option nearest the deficient option. The parameters of the model can also be adjusted to predict reverse attraction or *repulsion effects* (Spektor et al., 2018). It is also worth noting that DFT, LCA, AAM, and MLBA account for the related *reference point effects* (Tversky and Kahneman, 1991) using the same mechanisms used to account for the attraction effect. Horse race choice models, such as PR, are random utility models, and so inherit the triangle inequality and regularity properties, and cannot explain the attraction or reference point effects. The ADD model has never been applied to these effects either.

4.3 Compromise Effect

DFT's feedback matrix is again responsible for producing compromise effects. Because the compromise option is located between the extremes on both attribute dimensions,

its activation simultaneously suppresses activation in both extreme options, increasing relative preference for the compromise. LCA again uses loss aversion to account for compromise effects because the compromise option is always seen as having a small disadvantage relative to its competitor, which each have one small disadvantage and one large disadvantage. AAM and MLBA again use preprocessing to produce compromise effects. In the AAM, associations produce weights that favor the stronger attribute for the compromise option. MLBA primarily uses its subjective value function to produce extremeness aversion favoring the compromise. Neither the PR nor ADD models has been applied to this effect.

Usher et al. (2008) further investigated the dynamics of the compromise effect by first presenting a trinary choice set, but after participants made a choice, they were told that the chosen option was no longer available, and they would need to select again. Usher et al. (2008) found that when one extreme was initially selected, the compromise was most likely to be chosen next. Although this finding would appear to go against DFT's predictions that extreme options have correlated activation – because the other extreme was not chosen – Hotaling et al. (2010) argued that participants who show an initial preference for one extreme likely place greater weight on the attribute that option maximizes. In such cases, DFT predicts that the compromise will be chosen next because it now has the greatest value on this dimension.

4.4 Context Effects Under Time Pressure

The similarity effect has been found to increase with deliberation time in an inference task (Trueblood et al., 2014). MLBA predicts this effect, while MDFT, LCA, and AAM do not. Likewise, increased time pressure has been shown to decrease the compromise and attraction effects in consumer choice tasks (Pettibone, 2012), as well as reference point effects (Ashby et al., 2012). DFT produces this effect because the lateral inhibition represented in its feedback matrix builds over time. A similar story is true for LCA and AAM, whose loss aversion and association mechanisms, respectively, have a compounding effect with each additional time step. According to MLBA, greater deliberation time reduces the influence of random noise and increases the advantage of options with greater mean drift rate, e.g., the attraction and compromise options.

4.5 Eye Movements

Because eye movements contain information about how decision makers direct their attention, they can provide insight into choice strategies. DFT, LCA, and MLBA posit that options are compared within an attribute, while PR, AAM, and ADD predict the opposite – processing across attributes within an option. Noguchi and Stewart (2014) found comparisons of options within an attribute to be the dominant eye-movement pattern. Using a procedure to track the sequence of eye movements they observed participants making pair-wise comparisons of options within one attribute first, before moving on to the next attribute. This fits well with the processing assumptions of DFT, LCA, and MLBA. However, Noguchi and Stewart (2014) also found that participants compared only two options at a time. This contrasts with DFT and LCA, which, strictly speaking, assume that all options are compared along an attribute simultaneously.

The ADD model goes beyond those above to make several accurate predictions about fixation-driven decision biases, eye movement patterns, and their relation to choice. For instance, it predicts the observed finding of a strong positive relationship between looking time for an alternative and preference. It also accounts for the observation that an option is more likely to be chosen if it was the last option looked at. More recently, Cohen et al. (2017) tested several models that incorporated gaze patterns into a sequential sampling framework, and found that the best version was one that only updated attribute values for currently viewed alternative.

4.6 Quantitative Comparisons

Several papers have used quantitative analyses to compare models. Although none of these include all the models described above, and findings often vary from study to study, they nonetheless provide insights into which models best account for the entirety of choice behavior observed in a study.

Some of these comparisons use aggregate choice data. In support of DFT, Berkowitsch et al. (2014) found that the model outperformed a multivariate probit or multinomial logit model in predicting trinary choices when context effects were present. Hancock et al. (2018) found that DFT better predicted preferences for travel routes compared to a multinomial logit model. Trueblood et al. (2014) found that the MLBA fit aggregate trinary choice data from perceptual and inferential tasks better than DFT.

However, when fit to individuals' preferential choices, Hotaling and Rieskamp (2019) found DFT and MLBA performed equally well, with each model showing strength and weaknesses. Rieskamp (2008) also had success using DFT to model individual choices, and found that it predicts choice probabilities for binary choices between gambles better than a probabilistic version of prospect theory or a stochastic version of the priority heuristic model. Similarly, Scheibehenne et al. (2009) found that DFT predicts binary choices between gambles better than the proportional difference model (Gonzalez-Vallejo, 2002). Cohen et al. (2017) found that DFT and MLBA both outperformed heuristic choice models, with MLBA showing a slight advantage over DFT in preferential choice. Bhatia (2017) found that the AAM outperformed a probabilistic version of prospect theory in two large experiments investigating reference point effects in binary choice. Fitting to individuals and using eye movements to inform changes in drift rates, Krajbich et al. (2010) and Krajbich and Rangel (2011) showed that ADD predicts choices and response times better than a standard diffusion model (Ratcliff, 1978).

Hancock et al. (2021) made a number of methodological additions to DFT and MLBA to facilitate their application to naturalistic traffic choice data. They found that both models outperformed typical random utility and random regret minimization models in estimation and out-of-sample prediction. Recent work has also used hierarchical modeling techniques to compare models. Turner et al. (2018) compared several models of preferential choice and found that DFT predicted aggregate data best, but AAM and LCA performed better than DFT and MLBA when using a hierarchical Bayesian analysis that incorporated individual differences. Similarly, interesting results come from a large analysis by Evans et al. (2019) comparing DFT, LCA, AAM, and MLBA in 12 different studies of context effects. They found that DFT and MLBA performed equally well, and better than LCA and AAM when considering choice probabilities only. However, when

choices and response times were analyzed simultaneously, additional differences emerged. All models did equally well with the attraction effect, and for the similarity and compromise effect they predicted response times equally well. However, MLBA had a significant advantage over other models in simultaneously accounting for choice proportions, which the authors attributed to its assumption of between-trial variability.

5 ADDITIONAL AND NEW DIRECTIONS

This section briefly reviews some additional important ideas for choice modeling coming from cognitive psychology and neuroscience.

5.1 Decision by Sampling

One limitation of the previously mentioned sequential sampling models is that they do not specify the details of the sampling process other than assuming a mean input vector μ and a covariance matrix Φ for the Brownian motion process. Another cognitive theory, called decision by sampling, provides more detailed mechanisms for the sequential evaluations entering the accumulation process (Stewart et al., 2006; Stewart and Simpson, 2008). Stewart et al. (2006) postulate that the subjective values of an attribute value are derived on-the-fly from comparisons with samples of other attribute values drawn from long-term memory. Furthermore, this long-term memory reflects the values experienced from real-world distributions. Stewart et al. (2006) showed that the shapes of utility functions, decision weighting functions, and delay discounting functions estimated in decision research can be explained by the real-world distributions of gains, losses, probabilities, and delays that people experience.

5.2 Decision Neuroscience

More recently, sequential sampling models of decision making have attracted the interest of decision neuroscientists (Fehr and Rangel, 2011). Neuroscientists have been examining the neural basis for decision making in the brains of Macaque monkeys using single cell recording techniques (Gold and Shadlen, 2001, 2002; Platt, 2002; Schall, 2003). Both the choice and the decision time of the monkeys were accurately predicted, trial-by-trial, based on where and when the neural activation crossed a threshold bound. Moreover, sequential sampling models have been fit to choice and response time data and then used to make a priori predictions for electrophysiological activation (Smith and Ratcliff, 2004). Although this research was based on single cell recording of saccadic eye movements from monkeys, converging evidence has been reported using cognitive tasks with humans. Researchers have recorded electrophysiological potentials from humans during a categorization task, called the lateralized readiness potential (Gratton et al., 1988). These potentials were recorded from the scalp above the premotor cortex that signals preparation for left or right hand movements to a cue. They found that choice and response times were determined by accumulation of the lateralized readiness potential to a threshold criterion.

More direct evidence for accumulation to threshold as the basis for perceptual decision making has been obtained from functional magnetic resonance imaging (fMRI; Liu and

Pleskac, 2011). Another imaging study by Gluth et al. (2012) explored the predictive accuracy of sequential sampling models for preferential choices. In their experiment the participants received sequential information about a stock company that could be bought or rejected. The results showed that a time-variant sequential sampling model using a decreasing rather than a fixed decision threshold was best in predicting the decision time and choice behavior. Furthermore, the option's value as assumed by the sequential sampling model correlated with the brain activity in the ventromedial prefrontal cortex, the right and left orbitofrontal cortex, and the ventral striatum. These results indicate that the brain accumulates samples of information by forming an updated value representation in dopaminergic areas including the ventromedial prefrontal cortex and the ventral striatum. The conclusion from this research is the basic idea that decisions in the brain are based on the accumulation of noisy activation until a threshold is reached, which forms the basis for the sequential sampling models in cognitive science.

6 GENERAL CONCLUSIONS

In this chapter we have reviewed some of the main behavioral findings in behavioral economics, consumer research, and psychology. The list of findings provides a basis for evaluating the different approaches to predict and explain human choice behavior. On the one hand we have shown that standard utility approaches, such as fixed and random utility models, have difficulties in explaining why people violate principles, such as strong stochastic transitivity, independence, or regularity. On the other hand, we have shown that more cognitive models of decision making that also aim to describe how preferential choices evolve dynamically over time can provide explanations for these findings. These models have been the result of the long and rich history of choice modeling in psychology. This work has led to more sophisticated models that have also increased the complexity of the models. The increase in theoretical complexity has been driven by two sources – one is striving to explain the highly context-sensitive nature of human preferential choice behavior; the second is the attempt to account for more details of choice behavior including decision time, eye movements, information search, and choice confidence (Pleskac and Busemeyer, 2010; Ratcliff and Starns, 2009; Van Zandt, 2000). Beyond that, decision neuroscientists have even begun to explain the neurological underpinnings of decision making using electrophysiological recordings and/or functional magnetic resonance imaging during choice.

The psychological history of choice modeling began over 60 years ago with Luce's choice model (Luce, 1959), which was essentially a context-independent theory. Next, there were developments of Thurstone's (1959) theory to account for the effects of similarity between alternatives on choice. Aspect-based theories by Restle (1961) and Tversky (1972a) were also devised to account for similarity effects.

However, these earlier models turned out to be inadequate to explain more complicated findings such as compromise and attraction effects. Dynamic models such as the sequential sampling models (DFT, LCA, AAM) are the first to be able to account for the full range of context effects in a coherent manner and predict how these effects change as a function of deliberation time. These models also add explanatory power by accounting for the strong effects of deliberation time on choice probability, and they gain empirical

testability by making new predictions regarding the relations between choice probability and choice response time. More recent models, such as the decision by sampling model, provide a new capability for deriving the evaluations entering the sequential sampling accumulation process on the basis of long-term memory retrievals, where the memory is built from experience with real-world distributions. When developing sophisticated models, it is of course also important to examine to what extent the increased complexity of the developed models can be justified by the increased goodness-of-fit in predicting human behavior. Only when the qualitative and quantitative increase in predictive accuracy of a model is large enough, can the increased complexity be justified.

Although the goals of psychologists and economist with respect to choice modeling overlap to some extent, there are also some clear differences. On the one hand, economists are often primarily interested in applying simple choice models to large samples of people with relatively few data per person, and the primary goal is to obtain efficient estimates of economically relevant parameters. For this goal, robust choice models are required that provide statistically efficient and computationally practical parameter estimates for further economic analysis. On the other hand, psychologists are primarily interested in explaining the cognitive processes that provide predictions across a wide range of puzzling findings using models that can simultaneously account for various measurements of choice including choice selection, decision time, confidence, and brain activation. For this goal, more complex models are required that provide explanatory coherence and the power to predict new findings. But these are only two extreme positions, and more often researchers are interested in both of these goals. For example, a marketing researcher may want to predict market share of a new product to be introduced to the market. The market share will depend on the context of the market; that is, the other products available. Given the huge costs involved in introducing a new product to the market (e.g., product development, marketing) the precision with which the future share of the product can be predicted becomes essential. In this situation, using a complex choice model that requires higher methodological effort to be estimated can pay off in terms of obtaining better predictions of people's behavior.

ACKNOWLEDGMENT

This work was supported by NIDA grant 5R01DA030551 to the second author and by a NSF research grant 100014_130149 to the third author.

NOTES

1. The Thurstone model has always been a theory of choice. Initially it was applied to psychophysics (Thurstone, 1927), but later (Thurstone, 1959) it was applied to values and preferences (Bock and Jones, 1968).
2. Tony Marley (personal communication, Nov. 9, 2011) pointed out that if $\Gamma = \mathbf{0}$ then this model becomes a special case of the horse race model described in Marley and Colonius (1991) and satisfies regularity.
3. Strictly speaking, the LCA is not an OU process because it imposes a lower threshold on the preferences so that they reflect off a bound at zero. This bound is used to make the model more similar to neural activations that must always be positive.

REFERENCES

- Aschenbrenner, K. M., Albert, D., and Schmalhofer, F. (1984). Stochastic choice heuristics. *Acta Psychologica*, 56, 153–166.
- Ashby, F. G. (2000). A stochastic version of general recognition theory. *Journal of Mathematical Psychology*, 44, 310–329.
- Ashby, N., Dickert, S., and Glöckner, A. (2012). Focusing on what you own: Biased information uptake due to ownership. *Judgment and Decision Making*, 7(3), 254–267.
- Ben Zur, H., and Bresnitz, S. J. (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica*, 47, 89–104.
- Berkowitsch, N. A., Scheibehenne, B., and Rieskamp, J. (2014). Rigorously testing multialternative decision field theory against random utility models. *Journal of Experimental Psychology: General*, 143(3), 1331–1348.
- Bhatia, S. (2013). Associations and the accumulation of preference. *Psychological Review*, 120(3), 522–543.
- Bhatia, S. (2017). Comparing theories of reference-dependent choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(9), 1490–1507.
- Birnbaum, M. H., and Schmidt, U. (2008). An experimental investigation of violations of transitivity in choice under uncertainty. *Journal of Risk and Uncertainty*, 37, 77–91.
- Bock, R. D., and Jones, L. V. (1968). *The Measurement and Prediction of Judgment and Choice*. Oxford: Holden-Day.
- Böckenholt, U. (2006). Thurstonian-based analyses: Past, present, and future utilities. *Psychometrika*, 71, 615–629.
- Brown, S., and Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review*, 112, 117–128.
- Brown, S. D., and Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, 57(3), 153–178.
- Busemeyer, J. R. (1985). Decision making under uncertainty: A comparison of simple scalability, fixed-sample, and sequential-sampling models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 538–564.
- Busemeyer, J. R., and Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, 100, 432–459.
- Chu, Y. P., and Chu, R. L. (1990). The subsidence of preference reversals in simplified and market-like experimental settings: A note. *American Economic Review*, 80, 902–911.
- Cohen, A. L., Kang, N., and Leise, T. L. (2017). Multi-attribute, multi-alternative models of choice: Choice, reaction time, and process tracing. *Cognitive Psychology*, 98, 45–72.
- Cox, G. E., and Shiffrin, R. M. (2017). A dynamic approach to recognition memory. *Psychological Review*, 124(6), 795–860.
- DeGroot, M. H. (1970). *Optimal Statistical Decisions*. New York: McGraw-Hill.
- Dhar, R., Nowlis, S. M., and Sherman, S. J. (2000). Trying hard or hardly trying: An analysis of context effects in choice. *Journal of Consumer Psychology*, 9, 189–200.
- Diederich, A. (2003a). Decision making under conflict: Decision time as a measure of conflict strength. *Psychological Science*, 10, 353–359.
- Diederich, A. (2003b). MDFT account of decision making under time pressure. *Psychonomic Bulletin & Review*, 10, 157–166.
- Dutilh, G., and Rieskamp, J. (2016). Comparing perceptual and preferential decision making. *Psychonomic Bulletin & Review*, 23(3), 723–737.
- Erev, I., and Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*, 112, 912–931.
- Evans, N. J., Holmes, W. R., and Trueblood, J. S. (2019). Response-time data provide critical constraints on dynamic models of multi-alternative, multi-attribute choice. *Psychonomic Bulletin & Review*, 26(3), 901–933.
- Fehr, E., and Rangel, A. (2011). Neuroeconomic foundations of economic choice: Recent advances. *Journal of Economic Perspectives*, 25, 3–30.

- Fiedler, S., and Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in Psychology*, 3, 335.
- Fontanesi, L., Gluth, S., Spektor, M. S., and Rieskamp, J. (2019). A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin & Review*, 26, 1099–1121.
- Glöckner, A., and Betsch, T. (2008). Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *Judgment and Decision Making*, 3, 215–228.
- Gluth, S., Hotaling, J. M., and Rieskamp, J. (2017). The attraction effect modulates reward prediction errors and intertemporal choices. *Journal of Neuroscience*, 37(2), 371–382.
- Gluth, S., Rieskamp, J., and Büchel, C. (2012). Deciding when to decide: Time-variant sequential sampling models explain the emergence of value-based decisions in the human brain. *Journal of Neuroscience*, 32, 10686–10698.
- Gold, J. I., and Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, 5, 10–16.
- Gold, J. I., and Shadlen, M. N. (2002). Banburismus and the brain decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*, 36, 299–308.
- Gonzalez, C., and Dutt, V. (2011). Instance based learning: Integrating decisions from experience in sampling and repeated choice experiments *Psychological Review*, 118, 523–551.
- Gonzalez-Vallejo, C. (2002). Making trade-offs: A probabilistic and context-sensitive model of choice behavior. *Psychological Review*, 109, 137–154.
- Gratton, G., Coles, M. G. H., Sirevaag, E. J., Eriksen, C. W., and Donchin, E. (1988). Pre- and poststimulus activation of response channels: A psychophysiological analysis. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 331–344.
- Guo, F. Y., and Holyoak, K. J. (2002). Understanding similarity in choice behavior: A connectionist model. Paper presented at the Annual Meeting of the Cognitive Science Society, Mahwah, NJ.
- Hancock, T. O., Hess, S., and Choudhury, C. F. (2018). Decision field theory: Improvements to current methodology and comparisons with standard choice modelling techniques. *Transportation Research Part B: Methodological*, 107, 18–40.
- Hancock, T. O., Hess, S., Marley, A. A., and Choudhury, C. F. (2021). An accumulation of preference: Two alternative dynamic models for understanding transport choices. *Transportation Research Part B: Methodological*, 149, 250–282.
- Heath, T. B., and Chatterjee, S. (1995). Asymmetric decoy effects on lower quality versus higher quality brands: Meta-analytic and experimental evidence. *Journal of Consumer Research*, 22, 268–284.
- Hertwig, R., and Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences*, 13, 517–523.
- Hey, J. D. (2001). Does repetition improve consistency? *Experimental Economics*, 4, 5–54.
- Hotaling, J. M. (2020). Decision field theory-planning: A cognitive model of planning on the fly in multistage decision making. *Decision*, 7(1), 20–42.
- Hotaling, J. M., Busemeyer, J. R., and Li, J. (2010). Theoretical developments in decision field theory: A comment on K. Tsetsos, N. Chater, and M. Usher. *Psychological Review*, 117, 1294–1298.
- Hotaling, J. M., Cohen, A. L., Busemeyer, J. R., and Shiffrin, R. M. (2015). The dilution effect and information integration in perceptual decision making. *PLoS ONE*, 10(9), e0138481.
- Hotaling, J. M., and Rieskamp, J. (2019). A quantitative test of computational models of multialternative context effects. *Decision*, 6(3), 201–222.
- Huber, J., Payne, J. W., and Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9, 90–98.
- Huber, J., and Puto, C. (1983). Market boundaries and product choice: Illustrating attraction and substitution effects. *Journal of Consumer Research*, 10, 31–44.
- Iverson, G. J., and Falmagne, J.-C. (1985). Statistical issues in measurement. *Mathematical Social Sciences*, 10, 131–153.
- Janis, I. L., and Mann, L. (1977). *Decision Making: A Psychological Analysis of Choice, Conflict, and Decision Making*. New York: Free Press.
- Johnson, J. G., and Busemeyer, J. R. (2005). A dynamic, stochastic, computational model of preference reversal phenomena. *Psychological Review*, 112(4), 841–861.

- Kivetz, R., Netzer, O., and Srinivasan, V. (2004). Alternative models for capturing the compromise effect. *Journal of Marketing Research*, 41, 237–257.
- Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13, 1292–1298.
- Krajbich, I., Lu, D., Camerer, C., and Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology*, 3, 193.
- Krajbich, I., and Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, 108(33), 13852–13857.
- Laming, D. R. (1968). *Information Theory of Choice-Reaction Times*. New York: Academic Press.
- Lee, M. D., and Cummins, T. D. (2004). Evidence accumulation in decision making: Unifying the “take the best” and the “rational” models. *Psychonomic Bulletin & Review*, 11(2), 343–352.
- Link, S. W., and Health, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, 40, 77–111.
- Liu, T., and Pleskac, T. J. (2011). Neural correlates of evidence accumulation in a perceptual decision task. *Journal of Neurophysiology*, 106, 2383–2398.
- Loomes, G., Moffat, P. G., and Sugden, R. (2002). A microeconomic test of alternative stochastic theories of risky choice. *Journal of Risk and Uncertainty*, 24, 103–130.
- Lopes, L. L. (1996). When time is of the essence: Averaging, aspiration, and the short run. *Organizational Behavior & Human Decision Processes*, 65, 179–189.
- Louviere, J. J., Hensher, D. A., and Swait, J. D. (2000). *Stated Choice Models: An Analysis and Applications*. Cambridge: Cambridge University Press.
- Luce, R. D. (1959). *Individual Choice Behavior*. New York: Wiley.
- Luce, R. D. (2000). *Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches*. Mahwah, NJ: Lawrence Erlbaum.
- Luce, R. D., and Suppes, P. (1965). Preference, utility, and subjective probability. In R. D. Luce, R. B. Bush and E. Gallanter (eds.), *Handbook of Mathematical Psychology* (Vol. 3, pp. 249–410). New York: Wiley.
- Marley, A. A. J., and Colonius, H. (1991). The ‘horse race’ random utility model and its competing risks interpretation. *Journal of Mathematical Psychology*, 36, 1–20.
- Marschak, J. (1974). Binary-choice constraints and random utility indicators. In *Economic Information, Decision, and Prediction* (pp. 218–239). Dordrecht: Springer.
- May, K. (1954). Intransitivity, utility, and the aggregation of preference patterns. *Econometrica*, 22, 1–13.
- McFadden, D. (1981). Econometric models of probabilistic choice. In C. F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications* (pp. 198–272). Cambridge, MA: MIT Press.
- Mellers, B. A., and Biagini, K. (1994). Similarity and choice. *Psychological Review*, 101, 505–518.
- Mosteller, F., and Noguee, P. (1951). An experimental measurement of utility. *Journal of Political Economy*, 59, 371–404.
- Noguchi, T., and Stewart, N. (2014). In the attraction, compromise, and similarity effects, alternatives are repeatedly compared in pairs on single dimensions. *Cognition*, 132(1), 44–56.
- Nosofsky, R. M., and Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, 104, 266–300.
- Olschewski, S., and Rieskamp, J. (2021). Distinguishing three effects of time pressure on risk taking: Choice consistency, risk preference, and strategy selection. *Journal of Behavioral Decision Making*, 34(4), 541–554.
- Otter, T., Allenby, G. M., and van Zandt, T. (2008). An integrated model of discrete choice and response time. *Journal of Marketing Research*, 45, 593–607.
- Payne, J. W., Bettman, J. R., and Johnson, E. J. (1993). *The Adaptive Decision Maker*. New York: Cambridge University Press.
- Pettibone, J. (2012). Testing the effect of time pressure on asymmetric dominance and compromise in choice. *Judgment and Decision Making*, 7, 513–523.
- Platt, M. L. (2002). Neural correlates of decisions. *Current Opinion in Neurobiology*, 12, 141–148.

- Pleskac, T. J., and Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review*, 117, 864–901.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Ratcliff, R., and Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, 9(5), 347–356.
- Ratcliff, R., and Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology: Human Perception and Performance*, 26(1), 127–140.
- Ratcliff, R., and Smith, P. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111, 333–367.
- Ratcliff, R., and Starns, J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review*, 111, 333–367.
- Regenwetter, M., Dana, J., and Davis-Stober, C. P. (2011). Transitivity of preferences. *Psychological Review*, 118, 42–46.
- Restle, F. (1961). *Psychology of Judgment and Choice*. New York: Wiley.
- Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1446–1465.
- Rieskamp, J., Busemeyer, J. R., and Mellers, B. A. (2006). Extending the bounds of rationality: Evidence and theories of preferential choice. *Journal of Economic Literature*, 44, 631–661.
- Rieskamp, J., and Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*, 127, 258–276.
- Roe, R. M., Busemeyer, J. R., and Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, 108, 370–392.
- Rumelhart, D. E., and Greeno, J. G. (1971). Similarity between stimuli: An experimental test of the Luce and Restle choice models. *Journal of Mathematical Psychology*, 8, 370–381.
- Schall, J. D. (2003). Neural correlates of decision processes: Neural and mental chronometry. *Current Opinion in Neurobiology*, 13, 182–186.
- Scheibehenne, B., Rieskamp, J., and Gonzalez-Vallejo, C. (2009). Cognitive models of choice: Comparing decision field theory to the proportional difference model. *Cognitive Science*, 33, 911–939.
- Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, 16, 158–174.
- Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychological Review*, 102, 567–593.
- Smith, P. L., and Ratcliff, R. (2004). The psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27, 161–168.
- Spektor, M. S., Kellen, D., and Hotaling, J. M. (2018). When the good looks bad: An experimental exploration of the repulsion effect. *Psychological Science*, 29(8), 1309–1320.
- Stewart, N., Chater, N., and Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53, 1–26.
- Stewart, N., and Simpson, K. (2008). A decision-by-sampling account of decision under risk. In N. Chater and M. Oaksford (eds.), *The Probabilistic Mind: Prospects for Bayesian Cognitive Science* (pp. 261–276). Oxford: Oxford University Press.
- Svenson, O., and Edland, A. (1987). Changes of preference under time pressure: Choices and judgments. *Scandinavian Journal of Psychology*, 28, 322–330.
- Svenson, O., and Maule, A. J. (1993). *Time Pressures and Stress in Judgment and Decision Making*. New York: Plenum.
- Thurstone, L. L. (1927). Law of comparative judgment. *Psychological Review*, 34, 273–276.
- Thurstone, L. L. (1959). *The Measurement of Values*. Chicago: University of Chicago Press.
- Townsend, J. T., and Ashby, F. G. (1983). *Stochastic Modeling of Elementary Psychological Processes*. Cambridge: Cambridge University Press.
- Train, K. E. (2003). *Discrete Choice Methods with Simulation*. Cambridge: Cambridge University Press.
- Trueblood, J. S., Brown, S. D., and Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review*, 121(2), 179–205.

- Tsai, R. C., and Böckenholt, U. (2006). Modelling intransitive preferences: A random-effects approach. *Journal of Mathematical Psychology*, 50, 1–14.
- Tsetso, K., Usher, M., and Chater, N. (2010). Preference reversal in multialternative choice. *Psychological Review*, 117, 1275–1293.
- Turner, B. M., Schley, D. R., Muller, C., and Tsetso, K. (2018). Competing theories of multialternative, multiattribute preferential choice. *Psychological Review*, 125(3), 329–362.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76, 31–48.
- Tversky, A. (1972a). Elimination by aspects: A theory of choice. *Psychological Review*, 76, 281–299.
- Tversky, A. (1972b). Choice by elimination. *Journal of Mathematical Psychology*, 9, 341–367.
- Tversky, A., and Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106, 1039–1061.
- Tversky, A., and Russo, J. E. (1969). Similarity and substitutability in binary choices. *Journal of Mathematical Psychology*, 6, 1–12.
- Tversky, A., and Sattath, S. (1979). Preference trees. *Psychological Review*, 86, 542–572.
- Tversky, A., and Simonson, I. (1993). Context dependent preferences. *Management Science*, 39, 1179–1189.
- Usher, M., Elhalal, A., and McClelland, J. L. (2008). The neurodynamics of choice, value-based decisions, and preference reversal. In N. Chater and M. Oaksford (eds.), *The Probabilistic Mind: Prospects for Bayesian Cognitive Science* (pp. 277–300). Oxford: Oxford University Press.
- Usher, M., and McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108, 550–592.
- Usher, M., and McClelland, J. L. (2004). Loss aversion and inhibition in dynamic models of multialternative choice. *Psychological Review*, 111, 757–769.
- Van Zandt, T. (2000). ROC curves and confidence judgments in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 582–600.
- Vickers, D. (1979). *Decision Processes in Visual Perception*. New York: Academic Press.
- Wallsten, T., and Barton, C. (1982). Processing probabilistic multidimensional information for decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 361–384.
- Wallsten, T., and Barton, C. (1992). Processing probabilistic multidimensional information for making decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 361–364.
- Wedell, D. H. (1991). Distinguishing among models of contextually induced preference reversals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 767–778.
- Willemsen, M. C., Böckenholt, U., and Johnson, E. J. (2011). Choice by value coding and value construction: Processes of loss aversion. *Journal of Experimental Psychology: General*, 140(3), 303–324.
- Wright, P. (1972). The harassed decision maker: Time pressures, distractions, and the use of evidence. *Journal of Applied Psychology*, 59, 555–561.
- Yellot, J. I. (1977). The relation between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double exponential distribution. *Journal of Mathematical Psychology*, 15, 109–144.