JOURNAL OF MATHEMATICAL PSYCHOLOGY 37, 446–471 (1993)

The Relationship between Independent Race Models and Luce's Choice Axiom

CLAUS BUNDESEN

Copenhagen University

Four theorems concerning the relationship between independent race models and the choice axiom of R. D. Luce (1959, *Individual choice behavior*, New York: Wiley) are developed. In race models, selection (choice) is determined by a parallel processing race between elements in the choice set. Under weak assumptions, the choice axiom is equivalent to the condition that the hazard functions of the processing times of the elements are mutually proportional functions of time (Theorem 1). When this condition is satisfied, kth choices are equivalent to first choices from the reduced choice set (i.e., the original choice set minus the previously selected elements) (Theorem 2). Theorems 3 and 4 are general limit theorems. Whether or not the hazard functions are mutually proportional, the choice axiom describes the asymptotic choice probabilities approached under uniform expansion of the choice sets (Theorem 3). Moreover, in the limit, kth-choice probabilities are identical to first-choice probabilities (Theorem 4).

INTRODUCTION

This article develops four theorems concerning the relationship between independent race models and Luce's choice axiom. Briefly, a race model of selection (choice) is a model in which selection is determined by a parallel processing race between elements in the choice set. The winner of the race is the subject's choice. If more than one element is chosen, the winner is the first choice, number two is the second choice, and so forth. Race models have been used to study subjective deadlines in simple reaction time (e.g., Kornblum, 1973; Ollman & Billington, 1972), speed-accuracy trade-offs (e.g., Meyer, Irwin, Osman, & Kounios, 1988), reactions to stop-signals (e.g., Logan & Cowan, 1984; Ollman, 1973; Osman, Kornblum, & Meyer, 1986), Stroop phenomena (e.g., Morton & Chambers, 1973), processing of redundant signals (e.g., Meijers & Eijkman, 1977; Miller, 1982; van der Heijden, La Heij, & Boer, 1983), and selection from multielement displays (Bundesen, 1987; Bundesen, Shibuya, & Larsen, 1985; Shibuya & Bundesen, 1988). They form the basis of a general theory of automaticity (Logan, 1988), a unified theory of visual recognition and attentional selection (Bundesen, 1990), and a new approach to a large class of random utility models (Marley, 1989).

The choice axiom (Axiom 1 of Luce, 1959; reformulated by Luce & Galanter, 1963; see also Clarke, 1957; Shepard, 1957) states a condition on the way choices from different choice sets are interrelated. It states, in essence, that if some elements

0022-2496/93 $5.00

are removed from a choice set, then the relative probabilities of choices among the remaining elements are preserved, provided that the reduced choice set contains at least one element that has a positive probability of being chosen from the original choice set. The choice axiom is simple and the resulting computations are easy. For this reason, in part, the axiom has been used very widely. Luce (1977) provides a survey of *The Choice Axiom after Twenty Years*. Later applications include detailed studies of visual letter confusion (e.g., Townsend & Ashby, 1982; Townsend & Landon, 1982), studies of visual identification and categorization of separable-dimension stimuli (e.g., Nosofsky, 1985, 1986, 1989), and studies of partial report (e.g., Bundesen, Pedersen, & Larsen, 1984).

The relationship between race models and the choice axiom has been considered before (see Bundesen, 1990; Bundesen *et al.*, 1985; Luce & Green, 1972; Marley, 1989; Marley & Colonius, 1992; Townsend & Landon, 1983; Vorberg, 1992a; for closely related work, see Luce & Suppes, 1965; Vorberg, 1992b; Yellott, 1977, 1980). In particular, Luce and Green (1972) noted that if $X_1, X_2, ..., X_n$ are mutually independent exponentially distributed random variables (processing times) with rate parameters $v_1, v_2, ..., v_n$, respectively, then the probability that X_1 is the smallest among the n random variables equals $v_1/(v_1+v_2+ \cdots +v_n)$, which implies that the choice axiom holds for selection based on picking the minimum (the winner of the race). The same implication has been derived from the weaker assumption that $X_1, X_2, ..., X_n$ are mutually independent random variables that have hazard functions that are all proportional to each other (cf. Bundesen, 1990, Footnote 4; Marley & Colonius, 1992; Vorberg, 1992a, 1992b). This result is restated and complemented (by considering the converse) in Theorem 1 of the present article.

The assumptions used in proving Theorems 1–4 are stated in the next section. They are fairly weak. Given these assumptions, Theorem 1 essentially states that the choice axiom is equivalent to the condition that the hazard functions of the processing times of the elements in the choice sets are mutually proportional functions of time. Theorem 2 states that if the hazard functions are all proportional to each other, then kth choices are equivalent to first choices in the following sense. If R is the set of all elements remaining after selection of the first $k-1$ elements from the choice set U, then the probability that element i is the kth to be selected from U is the same as the probability that element i would be the first choice in selection from a choice set consisting exclusively of the members of R.[1]

Theorems 3 and 4 are general limit theorems. Let U_n be a choice set created from U by, for every element i in U, adding $n-1$ elements that are identical to element i (cf. Yellott, 1977), and consider the probability of selecting a type i element from U_n. Theorem 3 essentially states that, irrespective of the distributions of the processing times of the elements in U, the choice axiom holds for U_n in the limit

[1] Basically the same result has been proved by Vorberg (1992a). A proof of a special case of Theorem 2 (the case in which all hazard functions are constant) can be found in Bundesen *et al.* (1985); see also Vorberg and Ulrich (1987).

as $n \to \infty$. Theorem 4 states that, in the limit as $n \to \infty$, kth choices are equivalent to first choices.

The development that follows should be self-contained. In some places this has meant re-proving results that are already available in the literature.

Reprint requests should be addressed to Claus Bundesen, Psychological Laboratory, Copenhagen University, Njalsgade 90, DK-2300 Copenhagen S., Denmark. Much of the work reported in this article (including the proof of the *if* part of Theorem 1 and the proofs of Theorems 3 and 4) has previously been described in oral presentation at the Inaugural Meeting of the European Society for Cognitive Psychology in Nijmegen, The Netherlands, September 9–12, 1985. The preparation of the final version of the article was supported by a grant from the International Human Frontier Science Program Organization. Thanks are due to Hans Colonius, Hans-Georg Geissler, A. A. J. Marley, Dirk Vorberg, and an anonymous reviewer for valuable comments on previous drafts.

PRESUPPOSITIONS

Choice Axiom

Let S and U be choice sets such that $S \subseteq U$, and let $i \in U$. Let $P_S(i)$ be the probability that the subject's (first) choice is element i when S is the choice set, and let $P_U(i)$ be the probability that the choice is i when U is the choice set. Also, let $P_U(S)$ be the probability that the choice is an element from S when U is the choice set, that is, $P_U(S) = \sum_{j \in S} P_U(j)$. Reformulating Luce (1959, p. 6; Luce & Galanter, 1963, p. 218) we say that *the choice axiom holds for U* iff

$$P_S(i) = P_U(i)/P_U(S)$$

for all i, S such that $i \in S \subseteq U$ and $P_U(S) > 0$. A simple consequence is

LEMMA 1. *If the choice axiom holds for U and if $S \subseteq U$ such that $P_U(S) > 0$, then the choice axiom holds for S.*

Following Yellott (1977), a choice set U_n is called an *nth-order uniform expansion* of a choice set U if U_n can be created from U by, for every element i in U, adding $n-1$ elements that are identical to element i. For $i \in U$, $P_U(i)$ denotes the probability that the subject's (first) choice is element i when U is the choice set, and $P_{U_n}(i)$ denotes the probability that the choice is element i when U_n is the choice set. Let $P^*_{U_n}(i)$ denote the probability that the choice is a type i element (i.e., element i or one of the $n-1$ copies of element i) when U_n is the choice set. Then (cf. Yellott, 1977):

LEMMA 2. *If the choice axiom holds for U_n, then $P^*_{U_n}(i) = P_U(i)$.*

Race Models

A *race model* for selection (choice) is a model in which selection is determined by a parallel processing race between elements in the choice set: The first element that finishes processing is the subject's choice. If more than one element is chosen, the first element that finishes processing is the subject's first choice, the second element that finishes processing is the second choice, and so forth. Unless otherwise specified, it is assumed that all elements in the choice set begin to be processed (start the race) at the same moment in time. If processing times for individual elements in the choice set are mutually independent random variables, the model is called an *independent* race model.

The class of independent race models to be considered is constrained by the following assumptions. Let U be a finite set of elements (choice objects). Then:

Assumption 1. For every $i \in U$ and every choice set S such that $i \in S$, the probability distribution of the processing time of i is the same when i is presented in S as when i is presented in U.

Comment. By Assumption 1, the processing time distributions of elements in U are independent of the size and constituency of the choice sets in which the elements are presented. The choice sets in question need not be (proper or improper) subsets of U. An independent race model that satisfies Assumption 1 is said to be *unlimited in processing capacity* (cf. Bundesen, 1987). Extensions to models with limited processing capacity are considered in the final section of the article.

Assumption 2. For every $i \in U$, the processing time of i is a random variable with probability distribution function $F_i(t)$ and probability density function $f_i(t)$ so that

$$F_i(t) = \int_{-\infty}^{t} f_i(x)\, dx.$$

Moreover, $F_i(0) = 0$.

Assumption 3. For every $i \in U$, $f_i(t)$ is continuous at all but a finite number of points.

Assumption 4. For every $i \in U$, let $h_i(t)$ be the hazard function of the processing time of i, that is, $h_i(t) = f_i(t)/[1 - F_i(t)]$. For all i, S such that $i \in S \subseteq U$, $h_i(t)/\sum_{j \in S} h_j(t)$ is piecewise monotone.

Comment. A function f of a real variable is said to be *piecewise monotone* (or, equivalently, *in general monotone*; cf. Hobson, 1927/1957, p. 351) iff any finite interval in which f is defined can be divided into a finite number of parts (intervals) in each of which f is monotone.

Equivalence Theorems

Lemma on First Choices

Lemma 3. *Consider the choices predicted by an independent race model that satisfies Assumptions 1–4 for processing times of elements in a finite set U. For all* $i \in U$, $P^*_{U_n}(i)$ *is a constant independent of n iff the ratio* $h_i(t)/\sum_{j \in U} h_j(t)$ *is a constant independent of t for every time t at which the ratio exists.*

Proof. See Appendix A.

Theorem on First Choices

Definition 1. For any functions $h_i(t)$ and $h_j(t)$, $h_i(t) \propto h_j(t)$ (read "$h_i(t)$ is proportional to $h_j(t)$") iff (i) there is a constant $c > 0$ such that $h_i(t) = ch_j(t)$ for every $t \in E$, or (ii) $h_i(t) = 0$ for every $t \in E$, or (iii) $h_j(t) = 0$ for every $t \in E$, where E is the set of points at which both $h_i(t)$ and $h_j(t)$ are defined.

Theorem 1. *Consider the choices predicted by an independent race model that satisfies Assumptions 1–4 for processing times of elements in a finite set U. The choice axiom holds for every uniform expansion of every subset of U iff* $h_i(t) \propto h_j(t)$ *for all* $i, j \in U$.

Proof. See Appendix B.

Theorem on Later Choices

Theorem 2. *Consider the choices predicted by an independent race model that satisfies Assumptions 1–3 for processing times of elements in a finite choice set U. If* $h_i(t) \propto h_j(t)$ *for all* $i, j \in U$, *then kth choices* $(k = 2, 3, 4, ...)$ *are related to first choices as follows. If R is the set of all elements remaining after selection of the first* $k-1$ *elements from U, then the conditional probability* $P^{(k)}_U(i \mid R)$ *that element i is the kth to be selected from U is the same as the probability* $P_R(i)$ *that element i would be the first choice in selection from a choice set consisting exclusively of the members of R.*

Proof. See Appendix C.

Remark. Luce (1959, p. 72) proposed a *ranking postulate* to account for rank ordering. Let U be a choice set consisting of n elements, and let $(i_1, i_2, ..., i_n)$ be the ranking of U in which $i_1 \in U$ appears first, $i_2 \in U$ appears second, ..., and $i_n \in U$ appears last. The ranking postulate states that the probability of $(i_1, i_2, ..., i_n)$ equals

$$P_U(i_1)\, P_{U-\{i_1\}}(i_2) \cdots P_{U-\{i_1, i_2, ..., i_{n-1}\}}(i_n).$$

Theorems 1 and 2 imply a simple relationship between the choice axiom and the ranking postulate. Consider the choices predicted by an independent race model that satisfies Assumptions 1–4 for processing times of elements in a finite set U. If the choice axiom holds for every uniform expansion of every subset of U, then the ranking postulate holds for rankings of U by order of processing.

LIMIT THEOREMS

Theorem on First Choices

THEOREM 3. *Consider the choices predicted by an independent race model that satisfies Assumptions 1–4 for processing times of elements in a finite set U. For all i, S such that* $i \in S \subseteq U$, *if* $\lim_{n \to \infty} P^*_{U_n}(S) > 0$, *then*

$$\lim_{n \to \infty} P^*_{S_n}(i) = \lim_{n \to \infty} P^*_{U_n}(i) / \lim_{n \to \infty} P^*_{U_n}(S).$$

That is, the choice axiom describes the asymptotic choice probabilities approached under uniform expansion of U and its subsets.

Proof. See Appendix D.

Theorem on Later Choices

THEOREM 4. *Consider choices from* U_n *predicted by an independent race model that satisfies Assumptions 1–4 for processing times of elements in a finite set U. In the limit as* $n \to \infty$, *the probability that the kth element* ($k = 2, 3, 4, ...$) *selected from* U_n *is a type i element is identical to the probability that the first element selected from* U_n *is a type i element.*

Proof. See Appendix E.

EXTENSIONS AND EXAMPLES

Under weak assumptions (Assumptions 1–4), Theorem 1 states the conditions under which the choice axiom applies to the first-choice behavior of independent race models with unlimited processing capacity. Theorem 2 extends the analysis to kth-choice behavior. Theorems 3 (for first choices) and 4 (for later choices) lend a unique status to the choice axiom; for almost every independent race model with unlimited processing capacity, the choice axiom describes the asymptotic behavior approached under uniform expansion of the choice sets. In the present section, the analysis is extended to independent race models with limited processing capacity.

For some applications, Assumption 1 is too strong. For example, in many studies of visual selection from multielement displays (see Bundesen, 1990, for a review), the time taken to process an individual element has been found to be an increasing function of the number of elements in the display (i.e., the size of the choice set). Such findings suggest that elements are sampled from visual displays by a process with limited capacity.

To analyze independent race models with *limited processing capacity*, a quantitative notion of processing capacity is needed. Intuitively, variations in the amount of processing capacity allocated to an element concern the rate at which the element is processed, but not the type of processing that is done. More specifically Bundesen (1987) proposed that the effect of processing an element from time 0 to time t with a capacity of k units should equal the effect of processing the element from time 0 to time kt with a capacity of 1 unit. Thus, if $F_i(t)$ is the conditional distribution function for the processing time of element i given that the capacity allocated to element i is k units and $G_i(t)$ is the conditional distribution function for the processing time of element i given that the capacity allocated to element i is 1 unit, then for any time t,

$$F_i(t) = G_i(kt). \qquad (1)$$

The quantitative notion of capacity expressed in (1) is closely similar to the notion of capacity found in the multicomponent model of Rumelhart (1970).

Rumelhart (1970) introduced a notion of *attentional weights* to describe the way capacity is distributed among the elements in the choice set: For any elements i and j, there are attentional weights w_i and w_j such that the ratio between the amount of capacity allocated to element i and the amount of capacity allocated to element j equals the ratio between w_i and w_j. The amount of capacity allocated to a particular element is supposed to depend upon the choice set in which it is presented, but the attentional weight of the element is supposed to be constant across choice sets.

Consider an independent race model with limited processing capacity and constant attentional weights (i.e., attentional weights w_i, w_j that are constant across choice sets). Let $F_i^{(S)}(t)$ and $F_i^{(U)}(t)$ be the probability distribution functions of the processing time of element i when i is presented in choice sets S and U, respectively. Let C_S be the total processing capacity when S is the choice set, and let C_U be the total processing capacity when U is the choice set. Then the processing capacity allocated to element i equals $C_S w_i / \sum_{j \in S} w_j$, when S is the choice set, and $C_U w_i / \sum_{j \in U} w_j$ when U is the choice set. By (1), therefore,

$$F_i^{(S)}(t) = G_i\left[C_S\left(w_i \Big/ \sum_{j \in S} w_j\right) t\right]$$

and

$$F_i^{(U)}(t) = G_i\left[C_U\left(w_i \Big/ \sum_{j \in U} w_j\right) t\right].$$

Accordingly,

$$F_i^{(S)}(t) = F_i^{(U)}(k_S t),$$

where

$$k_S = C_S \sum_{j \in U} w_j \bigg/ \left(C_U \sum_{j \in S} w_j \right)$$

is independent of i. Thus, an independent race model with limited processing capacity and constant attentional weights satisfies

Assumption A. For every $i \in U$ and every choice set S such that $i \in S$, let $F_i^{(U)}(t)$ be the probability distribution function of the processing time of i when i is presented in U, and let $F_i^{(S)}(t)$ be the probability distribution function of the processing time of i when i is presented in S. For every S, there is a constant $k_S > 0$ such that for all $i \in U \cap S$,

$$F_i^{(S)}(t) = F_i^{(U)}(k_S t).$$

Assumption A is a generalization of Assumption 1. (Assumption 1 is the special case in which $k_S = 1$ for every S.) Consider the selection probabilities predicted by an independent race model that satisfies Assumption A. Let the choice set S be a subset of U_n for some $n \geqslant 1$. When the available processing time is unlimited, the selection probabilities are independent of the value of k_S. A change in k_S changes the time scale of the race between the elements of S, but otherwise the race is unaffected. Thus, for any value of k_S, the selection probabilities are the same as they would have been if $k_S = 1$. Therefore, for any independent race model that satisfies Assumption A, there is an independent race model that satisfies Assumption 1 such that the two models predict the same selection probabilities when the choice set is a subset of U_n and the available processing time is unlimited. Accordingly, Theorems 1–4 can be used in analyzing all independent race models that satisfy Assumption A, including models with limited processing capacity and constant attentional weights.

EXAMPLES. Shibuya and Bundesen (1988) proposed a fixed-capacity independent race model (FIRM) with constant attentional weights for partial report from briefly exposed visual displays. The model assumes that, on each trial, a fixed processing capacity of C units is distributed over the elements in the stimulus display (the choice set). The attentional weight of a target, w_1, is greater than the attentional weight of a distractor, w_0, and the ratio w_0/w_1 is a main parameter in the model (parameter α). The processing time of an element in the display is exponentially distributed with a rate parameter which is equal to the amount of processing capacity allocated to that element. For a display with T targets and D distractors, the exponential processing rate parameters are $\mu = Cw_1/(Tw_1 + Dw_0) = C/(T + \alpha D)$ for every target and $\alpha\mu$ for every distractor.

FIRM violates Assumption 1, but satisfies Assumption A. Let S be the set of elements in a stimulus display. When viewing time is unlimited, the selection probabilities predicted by FIRM are independent of the total processing capacity distributed over the elements of S. A change in the total processing capacity changes the time scale of the race between the elements of S, but otherwise the race is unaffected. Thus, for every S, the selection probabilities predicted by FIRM are the same as the selection probabilities predicted by an unlimited-capacity independent race model in which the processing time of an element is exponentially distributed with a rate parameter that equals the attentional weight of the element. This unlimited-capacity model satisfies Assumptions 1–4. Because the processing time distributions are exponential, the unlimited-capacity model also satisfies the conditions on hazard functions in Lemma 3 and Theorems 1 and 2. By Lemma 3, therefore, selection probabilities predicted by FIRM are invariant under uniform expansion of the choice sets. By Theorem 1, selection by FIRM conforms to the choice axiom. And by Theorem 2, kth choices are equivalent to first choices from reduced choice sets.

Shibuya (1991) considered a generalization of FIRM: a fixed-capacity independent race model with constant attentional weights and gamma-distributed processing times. Like FIRM, the model (gFIRM) satisfies Assumption A. Thus, when the available processing time is unlimited, the selection probabilities predicted by gFIRM are identical to the selection probabilities predicted by an unlimited-capacity independent race model in which the processing time of an element is gamma-distributed with its original shape parameter but a rate parameter that equals the attentional weight of the element. This unlimited-capacity model satisfies Assumptions 1–4. When shape parameters differ from 1, and attentional weights are unequal, the unlimited-capacity model violates the conditions on hazard functions in Lemma 3 and Theorem 1. By Lemma 3, then, the selection probabilities predicted by gFIRM are not invariant under uniform expansion of the choice sets, and by Theorem 1, selection by gFIRM violates the choice axiom. Nevertheless, by Theorems 3 (for first choices) and 4 (for later choices), the choice axiom describes the asymptotic selection probabilities approached under uniform expansion of the choice sets.

Shibuya and Bundesen (1988) tested FIRM against observed probability distributions of the number of correctly reported targets for 12 combinations of the number of targets T and the number of distractors D at exposure durations ranging from 20 to 200 ms. The combinations of T and D included uniform expansions, that is, combinations with different values of T, but the same value of D/T. The fit was close. Shibuya (1991) fitted gFIRM to the same data. Estimated values for the shape parameter of the gamma distribution were very close to 1, so the best-fitting version of gFIRM was virtually identical to FIRM. Such results provide empirical motivation for the study of independent race models and the choice axiom.

APPENDIX A: PROOF OF LEMMA 3

Domain and Import of $h_i(t)/\sum_{j \in U} h_j(t)$

By Assumption 3, function $f_i(t)$ is discontinuous at not more than a finite number of points. Values of $f_i(t)$ at such points are immaterial, and I regard $f_i(t)$ and density and hazard functions defined in terms of $f_i(t)$ (notably, $h_i(t)$ and $f_U(t)$) as *undefined* at such points.

Let D_U denote $\{t \mid \forall i \in U: f_i(t)$ is defined and continuous; $\exists i \in U: f_i(t) > 0$; $\forall i \in U: F_i(t) < 1\}$. Clearly,

$$D_U = \left\{ t \,\middle|\, \sum_{i \in U} f_i(t) > 0; \forall i \in U: F_i(t) < 1 \right\}$$

$$= \left\{ t \,\middle|\, \sum_{i \in U} h_i(t) > 0 \right\},$$

and for all $i \in U$, D_U is the domain of $h_i(t)/\sum_{j \in U} h_j(t)$.

D_U is an open set. To see this, note that $\sum_{i \in U} f_i(t)$ is defined at all but a number of points forming a finite set, M. On every adjacent interval of M, $\sum_{i \in U} f_i(t)$ is continuous. In each of these intervals, therefore, the set of points at which $\sum_{i \in U} f_i(t) > 0$ is open (cf., e.g., Lipschutz, 1965, Chap. 4). Hence the union of these sets, $\{t \mid \sum_{i \in U} f_i(t) > 0\}$, is open. For all $i \in U$, $F_i(t)$ is continuous everywhere, so the set of points at which $F_i(t) < 1$ is open. Hence, since U is finite, the intersection $\{t \mid \forall i \in U: F_i(t) < 1\}$ is open. Accordingly, $D_U = \{t \mid \sum_{i \in U} f_i(t) > 0\} \cap \{t \mid \forall i \in U: F_i(t) < 1\}$ is open.

Being an open set of real numbers, D_U consists of a countable number of disjoint open intervals. Let (a, b) be one of these open intervals and, for brevity, let $g(t)$ stand for $h_i(t)/\sum_{j \in U} h_j(t)$. As proved below, $g(t)$ is continuous on (a, b), and $g(a+)$ (i.e., $\lim_{t \to a+} g(t)$) exists.

Proof. For all $j \in U$, both $f_j(t)$ and $F_j(t)$ are continuous on (a, b), so $g(t)$ is also continuous on (a, b). Let (a, b') be a finite part of (a, b). Because $g(t)$ is piecewise monotone (Assumption 4), (a, b') can be divided into a finite number of disjoint subintervals in each of which $g(t)$ is monotone. Let (a, b'') be the leftmost of these subintervals. Since $g(t)$ is both bounded (between 0 and 1) and monotone in (a, b''), $g(t)$ converges as $t \to a+$ (Weierstrass' principle of continuity), that is, $g(a+)$ exists. Q.E.D.

The import of $h_i(t)/\sum_{j \in U} h_j(t)$ is clarified by noting that, for every $t \in D_U$,

$$P_U(i \mid \tau = t) = h_i(t) \Big/ \sum_{j \in U} h_j(t), \tag{A1}$$

where $P_U(i \mid \tau = t)$ is the conditional probability that element i is selected from U given that the selection occurs at time t.

Proof. Let s be a point such that $s > t$, but s and t belong to the same open interval (a, b) in D_U. Let $P_U(i \mid t < \tau \leqslant s)$ denote the conditional probability that i is selected from U given that the selection occurs after time t but at or before time s. Then

$$P_U(i \mid t < \tau \leqslant s) = \frac{\int_t^s f_i(x) \prod_{j \in U - \{i\}} [1 - F_j(x)]\, dx}{\sum_{i \in U} \int_t^s f_i(x) \prod_{j \in U - \{i\}} [1 - F_j(x)]\, dx}.$$

Using l'Hospital's rule,

$$\begin{aligned} P_U(i \mid \tau = t) &= \lim_{s \to t} P_U(i \mid t < \tau \leqslant s) \\ &= \lim_{s \to t} \frac{f_i(s) \prod_{j \in U - \{i\}} [1 - F_j(s)]}{\sum_{i \in U} f_i(s) \prod_{j \in U - \{i\}} [1 - F_j(s)]} \\ &= \lim_{s \to t} \frac{f_i(s)/[1 - F_i(s)]}{\sum_{i \in U} f_i(s)/[1 - F_i(s)]} \\ &= h_i(t) \Big/ \sum_{j \in U} h_j(t). \end{aligned}$$

Q.E.D.

It follows readily that, for every $t \in D_U$,

$$P^*_{U_n}(i \mid \tau = t) = P_U(i \mid \tau = t). \tag{A2}$$

Proof. $P^*_{U_n}(i \mid \tau = t) = nh_i(t)/\sum_{j \in U} nh_j(t) = h_i(t)/\sum_{j \in U} h_j(t) = P_U(i \mid \tau = t)$.

Behavior of Latency Distributions under Uniform Expansion of U

Consider the time at which selection takes place. Its probability distribution function is

$$F_U(t) = 1 - \prod_{j \in U} [1 - F_j(t)], \tag{A3}$$

when U is the choice set, and

$$F_{U_n}(t) = 1 - [1 - F_U(t)]^n \tag{A4}$$

when U_n is the choice set. The corresponding density functions are

$$f_U(t) = \sum_{i \in U} f_i(t) \prod_{j \in U - \{i\}} [1 - F_j(t)] \tag{A5}$$

and

$$f_{U_n}(t) = nf_U(t)[1 - F_U(t)]^{n-1}. \tag{A6}$$

Functions f_U and f_{U_n} are defined (and continuous) everywhere except at points t at which there is an element $i \in U$ such that $f_i(t)$ is undefined. We prove that, for every n,

$$D_U = \{t \mid f_{U_n}(t) > 0\}. \tag{A7}$$

Proof. The proof has three steps. Step 1 shows that, for every t,

$$f_U(t) > 0 \Rightarrow F_U(t) < 1, \tag{i}$$

and, for all $i \in U$,

$$f_i(t) > 0 \Rightarrow F_i(t) < 1. \tag{ii}$$

Suppose $F_U(t_0) = 1$. Then $F_U(t) = 1$ for every $t \geqslant t_0$, whence $\lim_{t \to t_0+} f_U(t) = 0$. Therefore, if f_U is continuous at t_0, then $f_U(t_0) = 0$. If f_U is not continuous at t_0, then f_U is undefined at t_0. In either case it is excluded that $f_U(t_0) > 0$, which establishes (i). Relation (ii) is established by similar reasoning.

Step 2 shows that, for every t, $f_{U_n}(t) > 0$ iff $f_U(t) > 0$. By (A6), if $f_{U_n}(t) > 0$, then $f_U(t) > 0$. Conversely, suppose $f_U(t) > 0$. By (i), $F_U(t) < 1$. Hence, by (A6), $f_{U_n}(t) > 0$.

Step 3 shows that, for every t, $f_U(t) > 0$ iff $t \in D_U$. Let $f_U(t) > 0$. By (A5), there is an element $i \in U$ such that $f_i(t) > 0$ and $F_j(t) < 1$ for all $j \in U - \{i\}$. By (ii), $F_i(t) < 1$, so for all $j \in U$, $F_j(t) < 1$. Hence, by the definition of D_U, $t \in D_U$. The converse follows immediately from (A5) and the definition of D_U. Steps 2 and 3 establish (A7). Q.E.D.

It follows that, for every n, $F_{U_n}(t)$ is strictly increasing in D_U. Indeed, for every $q \in D_U$, if $p < q$, then

$$F_{U_n}(p) < F_{U_n}(q). \tag{A8}$$

Proof. Let q belong to the open subinterval σ in D_U and let $p < q$. Then there is an $r \in \sigma$ such that $p < r < q$. Because the time derivative of F_{U_n}, f_{U_n}, is positive everywhere in σ, we get $F_{U_n}(r) < F_{U_n}(q)$. Because F_{U_n} is a probability distribution function, we have $F_{U_n}(p) \leqslant F_{U_n}(r)$. Hence $F_{U_n}(p) < F_{U_n}(q)$. Q.E.D.

Since D_U is an open set, an immediate consequence of (A8) is that, for all n and every $t \in D_U$,

$$0 < F_{U_n}(t) < 1. \tag{A9}$$

Finally we prove a strong result concerning the way in which the distribution of selection times changes under uniform expansion of the choice set. Consider selection from U_n as $n \to \infty$. If q is a time at which selection can occur, and $p < q$, then the probability that selection occurs after time q vanishes in relation to the probability that selection occurs in the interval from p up to q, however short this interval is. In symbols, for every $q \in D_U$, if $p < q$, then

$$\lim_{n \to \infty} \{[1 - F_{U_n}(q)]/[F_{U_n}(q) - F_{U_n}(p)]\} = 0. \tag{A10}$$

Proof. By (A8), the denominator is positive. By (A4), $[1 - F_{U_n}(q)]/[F_{U_n}(q) - F_{U_n}(p)]$ equals

$$[1 - F_U(q)]^n/\{[1 - F_U(p)]^n - [1 - F_U(q)]^n\},$$

which reduces [cf. (A9)] to

$$1/\{\{[1 - F_U(p)]/[1 - F_U(q)]\}^n - 1\}.$$

By (A8), this expression tends to zero as $n \to \infty$. Q.E.D.

Behavior of Choice Probabilities under Uniform Expansion of U

The probability of selecting a type i element, when U_n is the choice set, is given by

$$P^*_{U_n}(i) = \int_{D_U} f_{U_n}(t)\, P^*_{U_n}(i \mid \tau = t)\, dt$$

[cf. (A7)]. By (A1) and (A2), this reduces to

$$P^*_{U_n}(i) = \int_{D_U} f_{U_n}(t)\, g(t)\, dt, \tag{A11}$$

where $g(t)$ stands for $h_i(t)/\sum_{j \in U} h_j(t)$. Clearly, if $g(t)$ is a constant π_i for every $t \in D_U$, then regardless of n,

$$P^*_{U_n}(i) = \pi_i \int_{D_U} f_{U_n}(t)\, dt = \pi_i,$$

which proves the *if* part of Lemma 3.

The proof of the *only if* part of Lemma 3 is less direct. The first step is to prove that as $n \to \infty$, and selection from U_n becomes faster and faster, the probability of selecting a type i element from U_n approaches the asymptotic value of $h_i(t)/\sum_{j \in U} h_j(t)$ for $t \to \inf D_U$ (the greatest lower bound of D_U). That is,

$$\lim_{n \to \infty} P^*_{U_n}(i) = g(a_0 +), \tag{A12}$$

where a_0 stands for $\inf D_U$ and $g(a_0+) = \lim_{t \to a_0+} [h_i(t)/\sum_{j \in U} h_j(t)]$.

Proof. Consider $P^*_{U_n}(i \mid \tau \leqslant s)$, the conditional probability that the element selected from U_n is one of the type i elements given that the selection occurs at or before time s. For all n and every $s \in D_U$,

$$P^*_{U_n}(i \mid \tau \leqslant s) = \frac{\int_{t \leqslant s, t \in D_U} f_{U_n}(t)\, g(t)\, dt}{F_{U_n}(s)}.$$

Since

$$F_{U_n}(s) \inf_{t \leqslant s, t \in D_U} g(t) \leqslant \int_{t \leqslant s, t \in D_U} f_{U_n}(t)\, g(t)\, dt$$
$$\leqslant F_{U_n}(s) \sup_{t \leqslant s, t \in D_U} g(t),$$

it follows that

$$\inf_{t \leqslant s, t \in D_U} g(t) \leqslant P^*_{U_n}(i \mid \tau \leqslant s) \leqslant \sup_{t \leqslant s, t \in D_U} g(t).$$

By Assumption 2, D_U is bounded from below, so D_U has a greatest lower bound a_0. As $s \to a_0+$, both the infimum and the supremum of $g(t)$ in the interval $a_0 < t \leqslant s$ tend to $g(a_0+)$, so for every $\varepsilon > 0$, there is a time $s_0 > a_0$ such that, *regardless of n*,

$$|P^*_{U_n}(i \mid \tau \leqslant s_0) - g(a_0+)| < \varepsilon/2.$$

Clearly,

$$P^*_{U_n}(i) = P^*_{U_n}(i \mid \tau \leqslant s_0)\, F_{U_n}(s_0) + P^*_{U_n}(i \mid \tau > s_0)[1 - F_{U_n}(s_0)],$$

and, by (A4), $F_{U_n}(s_0) \to 1$ as $n \to \infty$, so $P^*_{U_n}(i)$ approaches $P^*_{U_n}(i \mid \tau \leqslant s_0)$ as $n \to \infty$. Hence there is a number N such that for all $n > N$,

$$|P^*_{U_n}(i) - P^*_{U_n}(i \mid \tau \leqslant s_0)| < \varepsilon/2$$

and therefore $|P^*_{U_n}(i) - g(a_0+)| < \varepsilon$. Q.E.D.

Now, assume that $P^*_{U_n}(i)$ is a constant independent of n. We prove that $g(t) = g(a_0+)$ for every $t \in D_U$, and we prove it indirectly. Remember that D_U consists of a countable number of disjoint open intervals. Number these intervals from left to right and let (a_0, b_0) be the first, (a_1, b_1) the second, ..., and (a_k, b_k) the $(k+1)$st. Suppose $g(t)$ is *not* constant at a value of $g(a_0+)$ in every one of the intervals. Let (a_k, b_k) be the leftmost interval in which $g(t)$ is not constant at a value of $g(a_0+)$. Four cases must be considered. In Case 1, $g(a_k+) > g(a_0+)$. In Case 2, $g(a_k+) < g(a_0+)$. In Cases 3 and 4, $g(a_k+) = g(a_0+)$. To distinguish between Cases 3 and 4, subdivide (a_k, b_k) into disjoint subintervals in each of which $g(t)$ is monotone (cf. Assumption 4) and consider the leftmost subinterval σ in which $g(t)$ is not constant. In Case 3, $g(t)$ is monotone nondecreasing in σ. In Case 4, $g(t)$ is monotone nonincreasing in σ.

Case 1. Let $c=[g(a_0+)+g(a_k+)]/2$. Then $g(a_0+)<c<g(a_k+)$. Since $g(t)$ is continuous on (a_k, b_k), there is a point q in (a_k, b_k) such that $g(t)>c$ for every t between a_k and q and, therefore,

$$\int_{a_k}^{q} f_{U_n}(t)\, g(t)\, dt > c[F_{U_n}(q)-F_{U_n}(a_k)].$$

For $t<a_k$, $g(t)$ is constant at a value of $g(a_0+)$, so

$$\int_{t<a_k,\, t\in D_U} f_{U_n}(t)\, g(t)\, dt = g(a_0+)\, F_{U_n}(a_k).$$

Since

$$P^*_{U_n}(i)=\int_{t<a_k,\, t\in D_U} f_{U_n}(t)\, g(t)\, dt+\int_{a_k}^{q} f_{U_n}(t)\, g(t)\, dt+\int_{t>q,\, t\in D_U} f_{U_n}(t)\, g(t)\, dt,$$

we get

$$P^*_{U_n}(i)>g(a_0+)\, F_{U_n}(a_k)+c[F_{U_n}(q)-F_{U_n}(a_k)], \tag{A13}$$

that is,

$$P^*_{U_n}(i)>g(a_0+)-g(a_0+)[1-F_{U_n}(a_k)]+c[F_{U_n}(q)-F_{U_n}(a_k)].$$

Accordingly, $P^*_{U_n}(i)>g(a_0+)$ if

$$g(a_0+)[1-F_{U_n}(a_k)]<c[F_{U_n}(q)-F_{U_n}(a_k)]. \tag{i}$$

Inequality (i) holds for all n if $g(a_0+)=0$. Suppose $g(a_0+)>0$. Then (i) is equivalent to

$$[1-F_{U_n}(q)]/[F_{U_n}(q)-F_{U_n}(a_k)]<[c-g(a_0+)]/g(a_0+). \tag{ii}$$

By (A10), there is a number N such that (ii) holds for all $n>N$. Thus, regardless of the value of $g(a_0+)$,

$$\text{there is an } N \text{ such that } P^*_{U_n}(i)>g(a_0+) \text{ for all } n>N. \tag{iii}$$

But the conjunction of (iii) and (A12) contradicts the assumption that $P^*_{U_n}(i)$ is a constant independent of n, which shows that Case 1 is excluded by this assumption.

Case 2. Let $c=[g(a_0+)+g(a_k+)]/2$. Then $g(a_0+)>c>g(a_k+)$. There is a point q in (a_k, b_k) such that $g(t)<c$ for every t between a_k and q and, therefore,

$$\int_{a_k}^{q} f_{U_n}(t)\, g(t)\, dt < c[F_{U_n}(q)-F_{U_n}(a_k)].$$

As in Case 1,

$$\int_{t<a_k,\, t\in D_U} f_{U_n}(t)\, g(t)\, dt = g(a_0+)\, F_{U_n}(a_k).$$

Because $g(t) \leqslant 1$,

$$\int_{t>q,\, t\in D_U} f_{U_n}(t)\, g(t)\, dt \leqslant 1 - F_{U_n}(q).$$

Hence,

$$P^*_{U_n}(i) < g(a_0+)\, F_{U_n}(a_k) + c[F_{U_n}(q) - F_{U_n}(a_k)] + [1 - F_{U_n}(q)], \qquad \text{(A14)}$$

that is,

$$P^*_{U_n}(i) < g(a_0+) - g(a_0+)[F_{U_n}(q) - F_{U_n}(a_k)] - g(a_0+)[1 - F_{U_n}(q)] + c[F_{U_n}(q) - F_{U_n}(a_k)] + [1 - F_{U_n}(q)].$$

Accordingly, $P^*_{U_n}(i) < g(a_0+)$ if

$$[c - g(a_0+)][F_{U_n}(q) - F_{U_n}(a_k)] + [1 - g(a_0+)][1 - F_{U_n}(q)] < 0. \qquad \text{(i)}$$

Inequality (i) holds for all n if $g(a_0+) = 1$. Suppose $g(a_0+) < 1$. Then (i) is equivalent to

$$[1 - F_{U_n}(q)]/[F_{U_n}(q) - F_{U_n}(a_k)] < [g(a_0+) - c]/[1 - g(a_0+)]. \qquad \text{(ii)}$$

By (A10), there is a number N such that (ii) holds for all $n > N$. Thus, regardless of the value of $g(a_0+)$,

there is an N such that $P^*_{U_n}(i) < g(a_0+)$ for all $n > N$. (iii)

But the conjunction of (iii) and (A12) contradicts the assumption that $P^*_{U_n}(i)$ is a constant independent of n, which shows that Case 2 is excluded by this assumption.

Case 3. There are points p and q in σ such that $p < q$, $g(a_0+) < g(p) < g(q)$, and for all t, t' in D_U, $g(t) \leqslant g(t')$ if $t < t' < q$. Therefore,

$$\int_{t<p,\, t\in D_U} f_{U_n}(t)\, g(t)\, dt > g(a_0+)\, F_{U_n}(p)$$

and

$$\int_p^q f_{U_n}(t)\, g(t)\, dt > g(p)[F_{U_n}(q) - F_{U_n}(p)],$$

whence

$$P^*_{U_n}(i) > g(a_0+)\, F_{U_n}(p) + g(p)[F_{U_n}(q) - F_{U_n}(p)]. \tag{i}$$

Inequality (i) is analogous to (A13), so Case 3 can be excluded by a proof which is similar to that by which Case 1 was excluded.

Case 4. There are points p and q in σ such that $p < q$, $g(a_0+) > g(p) > g(q)$, and for all t, t' in D_U, $g(t) \geqslant g(t')$ if $t < t' < q$. Therefore,

$$\int_{t < p,\, t \in D_U} f_{U_n}(t)\, g(t)\, dt < g(a_0+)\, F_{U_n}(p),$$

$$\int_p^q f_{U_n}(t)\, g(t)\, dt < g(p)[F_{U_n}(q) - F_{U_n}(p)],$$

and

$$\int_{t > q,\, t \in D_U} f_{U_n}(t)\, g(t)\, dt \leqslant 1 - F_{U_n}(q),$$

whence

$$P^*_{U_n}(i) < g(a_0+)\, F_{U_n}(p) + g(p)[F_{U_n}(q) - F_{U_n}(p)] + [1 - F_{U_n}(q)]. \tag{i}$$

Inequality (i) is analogous to (A14), so Case 4 can be excluded by a proof which is similar to that by which Case 2 was excluded.

Having excluded Cases 1–4, we conclude that if $P^*_{U_n}(i)$ is a constant independent of n, then $g(t) = g(a_0+)$ for every $t \in D_U$. Lemma 3 is thus established.

APPENDIX B: PROOF OF THEOREM 1

Proof of "Only If" Part. Let $S = \{i, j\} \subseteq U$. By assumption, the choice axiom holds for every uniform expansion S_n of S. By Lemma 2, $P^*_{S_n}(i)$ and $P^*_{S_n}(j)$ are independent of n. Hence, by Lemma 3, $h_i(t)/[h_i(t) + h_j(t)]$ is a constant π_i and $h_j(t)/[h_i(t) + h_j(t)]$ is a constant π_j for every time t at which the two ratios are defined, that is, for every point in $D_S = \{t \mid \sum_{i \in S} h_i(t) > 0\}$. Either $\pi_i > 0$ or $\pi_j > 0$. Suppose that $\pi_i > 0$. Then there is a number $k \geqslant 0$ such that $\pi_j/\pi_i = k$ and, accordingly, $h_j(t)/h_i(t) = k$ for every $t \in D_S$. If $k > 0$, then for every $t \in D_S$, $h_i(t) = ch_j(t)$, where $c = 1/k > 0$; if $k = 0$, then $h_j(t) = 0$ for every $t \in D_S$. A similar argument can be made if $\pi_j > 0$. Thus, whether $\pi_i > 0$ or $\pi_j > 0$, either (i) there is a constant $c > 0$ such that $h_i(t) = ch_j(t)$ for every $t \in D_S$, or (ii) $h_i(t) = 0$ for every $t \in D_S$, or (iii) $h_j(t) = 0$ for every $t \in D_S$.

To complete the proof, let E be the set of points at which both $h_i(t)$ and $h_j(t)$ are defined, and consider $h_i(t)$ and $h_j(t)$ at points $t \in E - D_S$. At such points,

$h_i(t) = h_j(t) = 0$. Therefore, given that one of Conditions (i)–(iii) holds for every $t \in D_S$, the same condition holds for every $t \in E$. Q.E.D.

Proof of "If" Part. Assumption 4 can be dispensed with in this part of the proof. Consider the choices predicted by an independent race model that satisfies Assumptions 1–3 for processing times of elements in U, and assume that $h_i(t) \propto h_j(t)$ for all $i, j \in U$. We first show that the choice axiom holds for U.

Let $S \subseteq U$ such that $P_U(S) > 0$. Then there is an element k in S such that $P_U(k) > 0$. As before, let $D_U = \{t \mid \sum_{i \in U} h_i(t) > 0\}$. Then $h_k(t) > 0$ for every $t \in D_U$. (To see this, suppose $h_k(t_0) = 0$ for some $t_0 \in D_U$. By the definition of D_U, there is an element $i \in U$ such that $h_i(t_0) > 0$. But $h_k(t_0) = 0$, $h_i(t_0) > 0$, and $h_k(t) \propto h_i(t)$ implies that $h_k(t) = 0$ for every $t \in D_U$, which contradicts the supposition that $P_U(k) > 0$.) Moreover, for all $i \in U$, there is a nonnegative constant v_i such that $h_i(t)/h_k(t) = v_i$ for every $t \in D_U$. Accordingly,

$$P_U(i) = \int_{D_U} f_U(t) \frac{h_i(t)}{\sum_{m \in U} h_m(t)}\, dt = \int_{D_U} f_U(t) \frac{h_i(t)/h_k(t)}{\sum_{m \in U} h_m(t)/h_k(t)}\, dt$$

$$= \int_{D_U} f_U(t)\, v_i \Big/ \sum_{m \in U} v_m \, dt = v_i \Big/ \sum_{m \in U} v_m, \tag{B1}$$

where $f_U(t)$ is the probability density function for the time of selection when U is the choice set.

Let $f_S(t)$ be the probability density function for the time of selection when S is the choice set, and let $D_S = \{t \mid \sum_{i \in S} h_i(t) > 0\}$. Because $S \subseteq U$, $\sum_{i \in S} h_i(t)$ is defined if $t \in D_U$. Moreover, since there is an element k in S such that $h_k(t) > 0$ for every $t \in D_U$, $\sum_{i \in S} h_i(t) > 0$ if $t \in D_U$. Thus, $D_U \subseteq D_S$. Conversely, if $t \in D_S$ (i.e., $\sum_{i \in S} h_i(t) > 0$) and $\sum_{i \in U} h_i(t)$ is defined, then $t \in D_U$ (i.e., $\sum_{i \in U} h_i(t) > 0$), so $D_S - D_U$ is the set of points t at which $h_i(t)$ is undefined for some $i \in U - S$. In fact, $D_S - D_U$ is the (finite) set of points at which $f_i(t)$ is undefined for some $i \in U - S$. To show this, we prove that for all $i \in U$, $F_i(t) < 1$ for every $t \in D_S$.

Proof. Clearly, for all $i \in U$, $F_i(t) < 1$ for every time t unless D_U is bounded from above. If D_U is bounded from above, then $F_i(t) < 1$ for every $t < \sup D_U$, and there is an element m in U such that

$$F_m(t) \to 1 \quad \text{as} \quad t \to \sup D_U. \tag{B2}$$

Note that $h_m(t) > 0$ for some $t \in D_U$. For every element j, the distribution and the hazard function are related by

$$F_j(t) = 1 - \exp\left(-\int_{-\infty}^{t} h_j(x)\, dx\right)$$

(cf., e.g., Luce, 1986, p. 15), so for all $j \in U$,

$$F_j(t) \to 1 \text{ as } t \to \sup D_U \Leftrightarrow \int_{-\infty}^{t} h_j(x)\,dx \to \infty \text{ as } t \to \sup D_U. \tag{B3}$$

As noted before, there is an element k in S such that $h_k(t) > 0$ for every $t \in D_U$. By assumption, $h_k(t) \propto h_m(t)$, whence there is a constant $c > 0$ such that $h_k(t) = c h_m(t)$ for almost every $t < \sup D_U$ (viz., for every point $t < \sup D_U$ such that $f_k(t)$ and $f_m(t)$ are both defined). Accordingly,

$$\int_{-\infty}^{t} h_k(x)\,dx \to \infty \text{ as } t \to \sup D_U \Leftrightarrow \int_{-\infty}^{t} h_m(x)\,dx \to \infty \text{ as } t \to \sup D_U. \tag{B4}$$

By (B3) and (B4),

$$F_k(t) \to 1 \text{ as } t \to \sup D_U \Leftrightarrow F_m(t) \to 1 \text{ as } t \to \sup D_U, \tag{B5}$$

and by (B2) and (B5),

$$F_k(t) \to 1 \qquad \text{as} \quad t \to \sup D_U,$$

so $\sup D_S \leqslant \sup D_U$. Hence, for all $i \in U$, $F_i(t) < 1$ for every $t \in D_S$. Q.E.D.

Since $D_U \subseteq D_S$, and $D_S - D_U$ is finite,

$$\begin{aligned} P_S(i) &= \int_{D_S} f_S(t)\, h_i(t) \Big/ \sum_{j \in S} h_j(t)\, dt \\ &= \int_{D_U} f_S(t) \frac{h_i(t)}{\sum_{j \in S} h_j(t)}\, dt = \int_{D_U} f_S(t) \frac{h_i(t)/h_k(t)}{\sum_{j \in S} h_j(t)/h_k(t)}\, dt \\ &= \int_{D_U} f_S(t)\, v_i \Big/ \sum_{j \in S} v_j\, dt = v_i \Big/ \sum_{j \in S} v_j. \end{aligned} \tag{B6}$$

By (B1) and (B6),

$$P_S(i) = \frac{v_i}{\sum_{j \in S} v_j} = \frac{v_i / \sum_{m \in U} v_m}{\sum_{j \in S} v_j / \sum_{m \in U} v_m} = \frac{P_U(i)}{P_U(S)},$$

so the choice axiom holds for U.

The above argument shows that the choice axiom holds for any finite set X if Assumptions 1–3 hold for X and if $h_i(t) \propto h_j(t)$ for all $i, j \in X$. Let V_n be an nth-order uniform expansion of some $V \subseteq U$. Given that Assumptions 1–3 hold for U and that $h_i(t) \propto h_j(t)$ for all $i, j \in U$, it follows that Assumptions 1–3 hold for V_n and that $h_i(t) \propto h_j(t)$ for all $i, j \in V_n$. By the above argument, therefore, the choice axiom holds for V_n. Q.E.D.

APPENDIX C: PROOF OF THEOREM 2

As Assumptions 1–3 hold for U, and $R \subseteq U$, Assumptions 1–3 also hold for R. Similarly, assuming that $h_i(t) \propto h_j(t)$ for all $i, j \in U$, we have $h_i(t) \propto h_j(t)$ for all $i, j \in R$. Accordingly, for all $i \in R$, there is a nonnegative constant v_i such that $h_i(t)/\sum_{j \in R} h_j(t) = v_i$ for every $t \in D_R$, where $D_R = \{t \mid \sum_{j \in R} h_j(t) > 0\}$. Hence, the probability that element i is the first choice in selection from a choice set consisting exclusively of the members of R is

$$P_R(i) = \int_{D_R} f_R(t)\, h_i(t) \Big/ \sum_{j \in R} h_j(t)\, dt = v_i,$$

where $f_R(t)$ is the probability density function for the time of selection when R is the choice set.

Now consider the kth choice from U given a condition of the form

$$\begin{gathered}\text{1st choice was of element } i_1 \text{ at time } \tau_1,\\ \text{2nd choice was of element } i_2 \text{ at time } \tau_2,\\ \vdots\\ (k-1)\text{st choice was of element } i_{k-1} \text{ at time } \tau_{k-1},\end{gathered}$$

such that $R = U - \{i_1, i_2, ..., i_{k-1}\}$. Since processing times for different elements are stochastically independent, the conditional probability density and distribution functions for the processing time of any element $i \in U$ depend upon only i, R, and τ_{k-1}. For all $i \in R$, we have

$$f_i(t \mid R, \tau_{k-1}) = \begin{cases} 0 & \text{for} \quad t < \tau_{k-1} \\ f_i(t)/[1 - F_i(\tau_{k-1})] & \text{for} \quad t > \tau_{k-1}, \end{cases}$$

$$F_i(t \mid R, \tau_{k-1}) = \begin{cases} 0 & \text{for} \quad t < \tau_{k-1} \\ [F_i(t) - F_i(\tau_{k-1})]/[1 - F_i(\tau_{k-1})] & \text{for} \quad t > \tau_{k-1}, \end{cases}$$

and, therefore,

$$h_i(t \mid R, \tau_{k-1}) = \frac{f_i(t \mid R, \tau_{k-1})}{1 - F_i(t \mid R, \tau_{k-1})} = \begin{cases} 0 & \text{for} \quad t < \tau_{k-1} \\ h_i(t) & \text{for} \quad t > \tau_{k-1}. \end{cases}$$

Similarly, the conditional probability density function for the time at which the kth selection from U occurs (i.e., the conditional probability density of the minimum completion time for the elements in R) is given by

$$f_R(t \mid \tau_{k-1}) = \begin{cases} 0 & \text{for} \quad t < \tau_{k-1} \\ f_R(t)/[1 - F_R(\tau_{k-1})] & \text{for} \quad t > \tau_{k-1}, \end{cases}$$

where f_R and F_R are the corresponding unconditional probability density and distribution functions. Let $M = \{t \mid t > \tau_{k-1},\ t \in D_R\}$. For all $i \in R$, then, the conditional probability that element i is the kth to be selected from U is

$$P_U^{(k)}(i \mid R, \tau_{k-1}) = \int_M f_R(t \mid \tau_{k-1})\, h_i(t \mid R, \tau_{k-1}) \Big/ \sum_{j \in R} h_j(t \mid R, \tau_{k-1})\, dt$$

$$= \int_M \frac{f_R(t)}{1 - F_R(\tau_{k-1})} \frac{h_i(t)}{\sum_{j \in R} h_j(t)}\, dt = v_i.$$

Thus, $P_U^{(k)}(i \mid R, \tau_{k-1})$ is independent of τ_{k-1}, and

$$P_U^{(k)}(i \mid R) = P_U^{(k)}(i \mid R, \tau_{k-1}) = P_R(i). \qquad \text{Q.E.D.}$$

APPENDIX D: PROOF OF THEOREM 3

For brevity, let $P^*_{U_\infty}(i)$ stand for $\lim_{n \to \infty} P^*_{U_n}(i)$. By Eq. (A12),

$$P^*_{U_\infty}(i) = \lim_{t \to a_0+} \left[h_i(t) \Big/ \sum_{j \in U} h_j(t) \right], \tag{D1}$$

where a_0 stands for $\inf D_U$. Given that Assumptions 1–4 hold for U, they also hold for any subset S of U. By analogy with (D1), therefore,

$$P^*_{S_\infty}(i) = \lim_{t \to \alpha_0+} \left[h_i(t) \Big/ \sum_{j \in S} h_j(t) \right], \tag{D2}$$

where α_0 stands for $\inf D_S$. We prove that

$$P^*_{S_\infty}(i) = \lim_{t \to a_0+} \left[h_i(t) \Big/ \sum_{j \in S} h_j(t) \right] \quad \text{if} \quad P^*_{U_\infty}(S) > 0, \tag{D3}$$

by showing that $\inf D_S = \inf D_U$ if $P^*_{U_\infty}(S) > 0$.

Proof. First, consider an arbitrary point $t_0 \leqslant \inf D_U$. For all $i \in U$, $F_i(t) = 0$ for every $t \leqslant t_0$, whence $\lim_{t \to t_0-} f_i(t) = 0$. Therefore, if f_i is continuous at t_0, then $f_i(t_0) = 0$, whence $h_i(t_0) = 0$; if f_i is not continuous at t_0, then f_i and h_i are undefined at t_0. Hence, $t_0 \notin D_S = \{t \mid \sum_{j \in S} h_j(t) > 0\}$. Thus, D_S contains no points at or below $\inf D_U$, that is, $\inf D_S \geqslant \inf D_U$.

Second, suppose $P^*_{U_\infty}(S) > 0$. Then there is an element k in S such that $P^*_{U_\infty}(k) > 0$ and, by (D1), $\lim_{t \to a_0+} [h_k(t) / \sum_{j \in U} h_j(t)] > 0$. Accordingly, as $h_k(t) / \sum_{j \in U} h_j(t)$ is continuous in some open interval (a_0, b_0), there is a $\delta > 0$ such that $h_k(t) / \sum_{j \in U} h_j(t) > 0$ for every $t \in (a_0, a_0 + \delta)$. Thus, $h_k(t) > 0$ and, therefore, $\sum_{j \in S} h_j(t) > 0$ for every $t \in (a_0, a_0 + \delta)$. Because $D_S = \{t \mid \sum_{j \in S} h_j(t) > 0\}$, the interval $(a_0, a_0 + \delta)$ is contained in D_S. Hence, $\inf D_S \leqslant a_0$, that is, $\inf D_S \leqslant \inf D_U$.

In conclusion, if $P^*_{U_\infty}(S) > 0$, then $\inf D_S = \inf D_U$, whence (D2) implies (D3). Q.E.D.

By (D1) and (D3), if $P^*_{U_\infty}(S) > 0$, then

$$\begin{aligned}\frac{P^*_{U_\infty}(i)}{P^*_{U_\infty}(S)} &= \frac{\lim[h_i(t)/\sum_{j\in U} h_j(t)]}{\lim \sum_{m\in S}[h_m(t)/\sum_{j\in U} h_j(t)]}\\ &= \lim \frac{h_i(t)/\sum_{j\in U} h_j(t)}{\sum_{m\in S}[h_m(t)/\sum_{j\in U} h_j(t)]}\\ &= \lim\left[h_i(t)\Big/\sum_{m\in S} h_m(t)\right] = P^*_{S_\infty}(i),\end{aligned}$$

where all limits are taken as $t \to a_0+$. Theorem 3 is thus established.

APPENDIX E: Proof of Theorem 4

Let $n \geqslant k$. By analogy with Eq. (A1), the conditional probability that the kth element selected from U_n is one of the type i elements, given that the kth selection occurs at time t and given that the $(k-1)$st selection occurred at time τ_{k-1}, where $\tau_{k-1} < t$, may be expressed as

$$P^{*(k)}_{U_n, R, \tau_{k-1}}(i \mid \tau_k = t) = \frac{n_i f_i(t \mid \tau_{k-1})/[1 - F_i(t \mid \tau_{k-1})]}{\sum_{j\in U}\{n_j f_j(t \mid \tau_{k-1})/[1 - F_j(t \mid \tau_{k-1})]\}},$$

where R is the set of all elements remaining after selection of the first $k-1$ elements, n_j is the number of type j elements in R,

$$f_j(t \mid \tau_{k-1}) = f_j(t)/[1 - F_j(\tau_{k-1})],$$

and

$$F_j(t \mid \tau_{k-1}) = [F_j(t) - F_j(\tau_{k-1})]/[1 - F_j(\tau_{k-1})].$$

Because

$$f_j(t \mid \tau_{k-1})/[1 - F_j(t \mid \tau_{k-1})] = f_j(t)/[1 - F_j(t)],$$

we get

$$P^{*(k)}_{U_n, R}(i \mid \tau_k = t) = n_i h_i(t)\Big/\sum_{j\in U} n_j h_j(t). \tag{E1}$$

For all $i \in U$,

$$n - k + 1 \leqslant n_i \leqslant n,$$

so

$$(n-k+1)\,h_i(t)\leqslant n_ih_i(t)\leqslant nh_i(t)$$

and

$$(n-k+1)\sum_{j\in U}h_j(t)\leqslant\sum_{j\in U}n_jh_j(t)\leqslant n\sum_{j\in U}h_j(t),$$

whence

$$\frac{n-k+1}{n}\frac{h_i(t)}{\sum_{j\in U}h_j(t)}\leqslant\frac{n_ih_i(t)}{\sum_{j\in U}n_jh_j(t)}\leqslant\frac{n}{n-k+1}\frac{h_i(t)}{\sum_{j\in U}h_j(t)},$$

and therefore

$$-\frac{k-1}{n}\frac{h_i(t)}{\sum_{j\in U}h_j(t)}\leqslant\frac{n_ih_i(t)}{\sum_{j\in U}n_jh_j(t)}-\frac{h_i(t)}{\sum_{j\in U}h_j(t)}$$
$$\leqslant\frac{k-1}{n-k+1}\frac{h_i(t)}{\sum_{j\in U}h_j(t)}.$$

Accordingly,

$$\left|\frac{n_ih_i(t)}{\sum_{j\in U}n_jh_j(t)}-\frac{h_i(t)}{\sum_{j\in U}h_j(t)}\right|\leqslant\frac{k-1}{n-k+1}\frac{h_i(t)}{\sum_{j\in U}h_j(t)}\leqslant\frac{k-1}{n-k+1}.$$

Thus, for every $\varepsilon>0$, there is a number n' such that for all $n>n'$ and every $t\in D_U$,

$$\left|\frac{n_ih_i(t)}{\sum_{j\in U}n_jh_j(t)}-\frac{h_i(t)}{\sum_{j\in U}h_j(t)}\right|<\frac{\varepsilon}{4}. \tag{E2}$$

As before, let $a_0=\inf D_U$. Then there is an open interval $(a_0, s_0)\subseteq D_U$ such that for every $t\in(a_0, s_0)$,

$$\left|\frac{h_i(t)}{\sum_{j\in U}h_j(t)}-g(a_0+)\right|<\frac{\varepsilon}{4}, \tag{E3}$$

where $g(a_0+)=\lim_{t\to a_0+}[h_i(t)/\sum_{j\in U}h_j(t)]$. By (E2) and (E3),

$$\left|\frac{n_ih_i(t)}{\sum_{j\in U}n_jh_j(t)}-g(a_0+)\right|<\frac{\varepsilon}{2} \tag{E4}$$

for all $n>n'$ and every $t\in(a_0, s_0)$. By (E1),

$$\inf_{t\in(a_0,s_0)}\left\{\frac{n_ih_i(t)}{\sum_{j\in U}n_jh_j(t)}\right\}\leqslant P^{*(k)}_{U_n,R}(i\mid\tau_k<s_0)\leqslant\sup_{t\in(a_0,s_0)}\left\{\frac{n_ih_i(t)}{\sum_{j\in U}n_jh_j(t)}\right\},$$

so (E4) implies that for all $n > n'$,

$$|P^{*(k)}_{U_n,R}(i \mid \tau_k < s_0) - g(a_0+)| \leqslant \varepsilon/2. \tag{E5}$$

As $n \to \infty$, the probability that the kth selection from U_n occurs before time s_0 tends to 1. Therefore, $P^{*(k)}_{U_n,R}(i)$ approaches $P^{*(k)}_{U_n,R}(i \mid \tau_k < s_0)$ as $n \to \infty$. That is, there is a number n'' such that for all $n > n''$,

$$|P^{*(k)}_{U_n,R}(i) - P^{*(k)}_{U_n,R}(i \mid \tau_k < s_0)| < \varepsilon/2. \tag{E6}$$

By (E5) and (E6), for all $n > \max(n', n'')$,

$$|P^{*(k)}_{U_n,R}(i) - g(a_0+)| < \varepsilon.$$

Thus, regardless of R and the value of k, the probability that the kth element selected from U_n is a type i element converges to $g(a_0+)$ as $n \to \infty$, which establishes Theorem 4.

References

Bundesen, C. (1987). Visual attention: Race models for selection from multielement displays. *Psychological Research*, **49**, 113–121.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, **97**, 523–547.

Bundesen, C., Pedersen, L. F., & Larsen, A. (1984). Measuring efficiency of selection from briefly exposed visual displays: A model for partial report. *Journal of Experimental Psychology: Human Perception and Performance*, **10**, 329–339.

Bundesen, C., Shibuya, H., & Larsen, A. (1985). Visual selection from multielement displays: A model for partial report. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 631–649). Hillsdale, NJ: Erlbaum.

Clarke, F. R. (1957). Constant-ratio rule for confusion matrices in speech communication. *Journal of the Acoustical Society of America*, **29**, 715–720.

Hobson, E. W. (1957). *The theory of functions of a real variable and the theory of Fourier's series* (Vol. 1, 3rd ed.). New York: Dover. (Original work published 1927)

Kornblum, S. (1973). Simple reaction time as a race between signal detection and time estimation: A paradigm and model. *Perception & Psychophysics*, **13**, 108–112.

Lipschutz, S. (1965). *General topology*. New York: Schaum.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492–527.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, **91**, 295–327.

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology*, **15**, 215–233.

Luce, R. D. (1986). *Response times. Their role in inferring elementary mental organization*. New York: Oxford Univ. Press.

Luce, R. D., & Galanter, E. (1963). Discrimination. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 191–243). New York: Wiley.

Luce, R. D., & Green, D. M. (1972). A neural timing theory for response times and the psychophysics of intensity. *Psychological Review*, **79**, 14–57.

LUCE, R. D., & SUPPES, P. (1965). Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 3, pp. 249–410). New York: Wiley.

MARLEY, A. A. J. (1989). A random utility family that includes many of the "classical" models and has closed form choice probabilities and choice reaction times. *British Journal of Mathematical and Statistical Psychology*, **42**, 13–36.

MARLEY, A. A. J., & COLONIUS, H. (1992). The "horse race" random utility model for choice probabilities and reaction times, and its competing risks interpretation. *Journal of Mathematical Psychology*, **36**, 1–20.

MEIJERS, L. M. M., & EIJKMAN, E. G. J. (1977). Distributions of simple RT with single and double stimuli. *Perception & Psychophysics*, **22**, 41–48.

MEYER, D. E., IRWIN, D. E., OSMAN, A. M., & KOUNIOS, J. (1988). The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition. *Psychological Review*, **95**, 183–237.

MILLER, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, **14**, 247–279.

MORTON, J., & CHAMBERS, S. M. (1973). Selective attention to words and colours. *Quarterly Journal of Experimental Psychology*, **25**, 387–397.

NOSOFSKY, R. M. (1985). Overall similarity and the identification of separable-dimension stimuli: A choice model analysis. *Perception & Psychophysics*, **38**, 415–432.

NOSOFSKY, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, **115**, 39–57.

NOSOFSKY, R. M. (1989). Further tests of an exemplar-similarity approach to relating identification and categorization. *Perception & Psychophysics*, **45**, 279–290.

OLLMAN, R. T. (1973). Simple reactions with random countermanding of the "go" signal. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 571–581). New York: Academic Press.

OLLMAN, R. T., & BILLINGTON, M. J. (1972). The deadline model for simple reaction times. *Cognitive Psychology*, **3**, 311–336.

OSMAN, A., KORNBLUM, S., & MEYER, D. E. (1986). The point of no return in choice reaction time: Controlled and ballistic stages of response preparation. *Journal of Experimental Psychology: Human Perception and Performance*, **12**, 243–258.

RUMELHART, D. E. (1970). A multicomponent theory of the perception of briefly exposed visual displays. *Journal of Mathematical Psychology*, **7**, 191–218.

SHEPARD, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, **22**, 325–345.

SHIBUYA, H. (1991). Comparison between stochastic models for visual selection. In J.-P. Doignon & J.-C. Falmagne (Eds.), *Mathematical psychology: Current developments* (pp. 337–356). New York: Springer-Verlag.

SHIBUYA, H., & BUNDESEN, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 591–600.

TOWNSEND, J. T., & ASHBY, F. G. (1982). Experimental test of contemporary mathematical models of visual letter recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **8**, 834–864.

TOWNSEND, J. T., & LANDON, D. E. (1982). An experimental and theoretical investigation of the constant-ratio rule and other models of visual letter confusion. *Journal of Mathematical Psychology*, **25**, 119–162.

TOWNSEND, J. T., & LANDON, D. E. (1983). Mathematical models of recognition and confusion in psychology. *Mathematical Social Sciences*, **4**, 25–71.

VAN DER HEIJDEN, A. H. C., LA HEIJ, W., & BOER, J. P. A. (1983). Parallel processing of redundant targets in simple visual search tasks. *Psychological Research*, **45**, 235–254.

VORBERG, D. (1992a). Within-stage independence, future-order independence, and realizability of parallel processing systems. Manuscript submitted for publication.

VORBERG, D. (1992b). Characterizations of independent random utility models which imply the choice axiom. Manuscript submitted for publication.

VORBERG, D., & ULRICH, R. (1987). Random search with unequal search rates: Serial and parallel generalizations of McGill's model. *Journal of Mathematical Psychology*, **31**, 1–23.

YELLOTT, J. I., JR. (1977). The relationship between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double exponential distribution. *Journal of Mathematical Psychology*, **15**, 109–144.

YELLOTT, J. I., JR. (1980). Generalized Thurstone models for ranking: Equivalence and reversibility. *Journal of Mathematical Psychology*, **22**, 48–69.

RECEIVED: October 29, 1990