

UC Irvine

UC Irvine Previously Published Works

Title

Generalized thurstone models for ranking: Equivalence and reversibility

Permalink

<https://escholarship.org/uc/item/9vs9m6w2>

Journal

Journal of Mathematical Psychology, 22(1)

ISSN

0022-2496

Author

Yellott, John I

Publication Date

1980-08-01

DOI

10.1016/0022-2496(80)90046-2

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Generalized Thurstone Models for Ranking: Equivalence and Reversibility¹

JOHN I. YELLOTT, JR.

*Cognitive Science Program, School of Social Sciences,
University of California, Irvine, California 92717*

Previous work has determined the conditions under which generalized versions of Thurstone's theory of comparative judgment are formally equivalent (i.e., empirically indistinguishable) for choice experiments. This note solves the analogous problem for ranking experiments: It is shown that if two "Generalized Thurstone Models" are equivalent for choice experiments with n alternatives they are also equivalent for ranking experiments with n alternatives, despite the fact that ranking generates many more preference probabilities. This result in turn allows one to determine which Generalized Thurstone Models are "reversible," i.e., satisfy the requirement that regardless of whether the subject ranks from best to worst or from worst to best, rankings that express the same preference order will occur with the same probability.

1. INTRODUCTION

1.1. *Equivalence*

Several recent articles in this journal have dealt with the equivalence properties of a family of random utility models for choice experiments that can be regarded as generalizations of Thurstone's (1927) Theory of Comparative Judgment—generalized in the sense that the utility random variables (Thurstone's "discriminal processes") are no longer required to have normal distributions. ("Equivalence" here means experimental indistinguishability. Two models—that is, theories—are said to be equivalent for a given class of experiments if results that satisfy one always satisfy the other—so that no experimental decision can be made between them. Yellott (1977, 1978), Moszner (1978), and Rockwell and Yellott (1979), all deal with a generalized version of Thurstone's Case V in which the utility random variables are independent and identically distributed except for shifts: Such models were referred to there simply as "Thurstone models"; here they will be called "Generalized Case V Thurstone (GT-V) Models." Strauss (1979), analyzes a broader class of "Generalized Thurstone (GT) Models" which includes the GT-V models and also nonindependent cases. All this work is reviewed here in Section 2.) This paper deals with the equivalence properties of these same models when they are applied to ranking experiments—that is, experiments in which the subject does not simply choose a single best alternative from some set, but instead rank orders all the alternatives from best to

¹ Research supported by the National Science Foundation, Grant No. BN577-27333.

worst (or vice versa). These properties can be quickly determined, and turn out to be both simple and somewhat counterintuitive. Because ranking experiments yield many more preference probabilities than choice experiments (i.e., $n!$ vs n), intuition suggests that they ought to be more efficient at discriminating between models. Specifically, one might expect to find at least some cases in which models that are nonidentical but nevertheless equivalent for certain choice experiments would become nonequivalent when applied to ranking versions of those same experiments. This can definitely happen within the general class of random utility models: Table 1 in Section 3 shows an example. Surprisingly, however, it proves to be impossible for GT models: Theorem 1 in Section 3.3 shows that two such models are equivalent for ranking experiments with n alternatives if and only if they are equivalent for choice experiments with n alternatives. In other words whenever two GT models are indistinguishable by choice experiments they are also indistinguishable by ranking experiments, despite the extra preference probabilities generated by the latter.

1.2. *Reversibility*

This equivalence theorem turns out to be useful in understanding an old puzzle in choice theory: The fact that Luce's (1959, 1977) Choice Axiom is incompatible with the intuitively sensible requirement that ranking probabilities should be essentially the same whether the subject ranks from best to worst or vice versa. In designing any ranking experiment one has to decide in advance on the direction of ranking, i.e., should the best alternative be assigned rank 1, the next-best rank 2, and so on ("best-to-worst" ranking), or should the worst alternative get rank 1, the next-worst rank 2, etc. ("worst-to-best" ranking)? In many contexts this decision seems entirely arbitrary because we expect that regardless of the direction instructions, rank orderings that mean the same thing ought to occur with the same probability. For example, if the alternatives are three tones a_1, a_2, a_3 to be ranked on the basis of loudness, it seems reasonable to expect that the probability of producing any given rank order, say $\langle a_i, a_j, a_k \rangle$, under the instruction "rank from loudest to softest" should be the same as the probability of producing the reverse ordering $\langle a_k, a_j, a_i \rangle$ under the instruction "rank from softest to loudest."

However, it has been recognized for a long time that this "reversibility" assumption is surprisingly difficult to reconcile with other theoretical notions that seem equally plausible—and that apparently have nothing to do with the direction of ranking (Luce, 1959; Block & Marschak, 1960; Luce & Suppes, 1965; Marley, 1968; Thorsen & Stever, 1974). In particular, it is well known that reversibility cannot easily be combined with the ideas that choice and ranking probabilities should satisfy a common random utility model and that ranking is accomplished by a series of independent choices—a pair of assumptions closely related to the Choice Axiom.

The difficulties here are neatly captured in a well-known impossibility theorem due to Block *et al.* (1960; Theorem 51 in Luce *et al.* 1965). The theorem holds for any number of alternatives but for convenience at this point we consider the simplest case, where there are only three. Suppose the alternatives are a_1, a_2, a_3 , and that both choice and ranking probabilities are experimentally determined for all subsets under best-to-

worst instructions and also under worst-to-best instructions. (For choices this means we sometimes ask the subject to choose the best alternative and sometimes to choose the worst.) Let $p(i)$ denote the probability of choosing a_i from $\{a_1, a_2, a_3\}$ under best-to-worst instructions; $p(i, j)$ the probability of choosing a_i from $\{a_i, a_j\}$ under the same instruction; $r(i, j, k)$ the probability of producing the rank order $\langle a_i, a_j, a_k \rangle$ in the best-to-worst condition when a_1, a_2 , and a_3 are all available; and $r(i, j)$ the probability of producing rank order $\langle a_i, a_j \rangle$ when only a_i and a_j are available. And let $p^*(i)$, $p^*(i, j)$, $r^*(i, j, k)$, and $r^*(i, j)$ denote the probabilities of the same events under worst-to-best instructions, i.e., $r^*(i, j, k)$ is the probability of saying that a_i is worst than a_j is worse than a_k , etc.

Now to capture the idea that both choices and rankings are generated by a common set of underlying utility random variables, we assume that the probability of choosing a_i over any set of competitors is the same as the probability that a_i is ranked ahead of those competitors, i.e.,

$$\begin{aligned} p(i) &= r(i, j, k) + r(i, k, j), \\ (i)^2 \quad p(i, j) &= r(i, j) = r(i, j, k) + r(i, k, j) + r(k, i, j), \\ p^*(i) &= r^*(i, j, k) + r^*(i, k, j), \\ p^*(i, j) &= r^*(i, j) = r^*(i, j, k) + r^*(i, k, j) + r^*(k, i, j). \end{aligned}$$

Next, to express the idea that ranking is accomplished by a series of independent choices, we assume that the probability of producing the rank order $\langle a_i, a_j, a_k \rangle$ is the same as the probability of first choosing a_i from the whole set $\{a_1, a_2, a_3\}$ and then, in an independent pairwise choice, choosing a_j over a_k :

$$\begin{aligned} (ii) \quad r(i, j, k) &= p(i) p(j, k), \\ r^*(i, j, k) &= p^*(i) p^*(j, k). \end{aligned}$$

(This assumption is sometimes referred to as "decomposition", e.g., in Strauss (1979).)

Finally, to express the notion that rank orderings that mean the same thing should have the same probability regardless of the instructed direction of ranking, we assume "reversibility":

$$\begin{aligned} (iii) \quad r(i, j) &= r^*(j, i), \\ r(i, j, k) &= r^*(k, j, i) \end{aligned}$$

Now taken one at a time each of these assumptions seems plausible, and even innocuous. However, Block *et al.* (1960), extending results obtained originally by Luce (1959),

² Assumption (i) is equivalent to the assumption (i') that there exist random variables U_1, U_2, U_3 such that for all i, j, k : $p(i) = P(U_i = \max\{U_1, U_2, U_3\})$; $p(i, j) = r(i, j) = P(U_i > U_j)$; and $r(i, j, k) = P(U_i > U_j > U_k)$. The fact that (i') \rightarrow (i) is obvious; the fact that (i) \rightarrow (i') is Block and Marschak's (1960) well-known characterization theorem for random utility models (Theorem 49 in Luce *et al.*, 1965).

showed that if all three hold the subject must be completely indifferent between the alternatives, i.e., (i), (ii), and (iii) together imply that $\forall i, j, k, p(i) = p^*(i) = \frac{1}{3}$; $p(i, j) = p^*(i, j) = \frac{1}{2}$, and $r(i, j, k) = r^*(i, j, k) = \frac{1}{6}$.

The proof of this remarkable fact requires only two simple steps. First, note that (i) and (ii) imply that both the best-to-worst and worst-to-best choice probabilities satisfy Luce's Choice Axiom, which for this three-alternative case is simply

$$(iv) \quad \begin{aligned} p(i, j) &= \frac{p(i)}{p(i) + p(j)}, \\ p^*(i, j) &= \frac{p^*(i)}{p^*(i) + p^*(j)}. \end{aligned}$$

(To show this use (ii) to substitute for $r(i, j, k)$ in the second part of (i) yielding

$$p(i, j) = p(i)p(j, k) + p(i)p(k, j) + p(k)p(i, j).$$

Then (iv) follows immediately. It is also easy to show that (i) and (iv) together imply (ii), so in the context of the random utility assumption (i), decomposition and the Choice Axiom are equivalent.) The Choice Axiom in turn implies the existence of ratio scale values v_1, v_2, v_3 and v_1^*, v_2^*, v_3^* such that

$$\begin{aligned} p(i, j) &= \frac{v_i}{v_i + v_j}; & p(i) &= \frac{v_i}{v_i + v_j + v_k}, \\ p^*(i, j) &= \frac{v_i^*}{v_i^* + v_j^*}; & p^*(i) &= \frac{v_i^*}{v_i^* + v_j^* + v_k^*}. \end{aligned}$$

The second step is to note that the first part of the reversibility assumption (iii) implies $v_i/v_j = v_i^*/v_j^*$. Since we can arbitrarily set $v_1 = v_1^* = 1$, it follows that in general $v_i^* = 1/v_i$, and then the second part of (iii) together with (ii) implies

$$\frac{v_i}{(v_i + v_j + v_k)} \frac{v_j}{(v_j + v_k)} = \frac{v_k^{-1}}{v_i^{-1} + v_j^{-1} + v_k^{-1}} \frac{v_j^{-1}}{v_i^{-1} + v_j^{-1}}.$$

Simplifying this equation we obtain $v_i v_k = v_j^2$, and first setting $k = 1$ (to yield $v_2 v_3 = 1$) and then $k = 2$ (to yield $v_1 v_3 = v_3 = v_2^2 = v_3^{-2}$, i.e., $v_3^3 = 1$) it follows that $v_1 = v_2 = v_3 = v_1^* = v_2^* = v_3^* = 1$.

Now this impossibility theorem is puzzling for two reasons. First, and foremost, it is not at all obvious why the Choice Axiom should have anything to do with the direction of ranking, and so the fact that it cannot be reconciled with reversibility seems to be nothing more than an algebraic accident. Second, the theorem does not give any hint as to the degree of incompatibility between reversibility and the Choice Axiom: We know that $r(i, j, k)$ and $r^*(k, j, i)$ cannot be equal (except in the degenerate case $v_i = v_i^* \equiv 1$),

but we have no intuitive basis for expecting their difference to be large or small. In fact a computer search shows that this difference is never greater than 0.05, and consequently would be practically impossible to detect experimentally. However it is not easy to see intuitively why this should be so.

Approached from the standpoint of GT models, however, the impossibility theorem becomes more understandable. The starting point is the fact that the Choice Axiom is equivalent to the GT-V model based on the double exponential distribution $P(\mathbf{X} \leq x) = e^{-e^{-x}}$ (Yellott, 1977). Consequently (i) and (ii) imply the simultaneous existence of real constants m_1, m_2, m_3 , and m_1^*, m_2^*, m_3^* , such that $\forall i, j, k$

$$\begin{aligned} p(i, j) &= P(m_i + \mathbf{X}_i > m_j + \mathbf{X}_j), \\ p(i) &= P(m_i + \mathbf{X}_i = \max\{m_j + \mathbf{X}_j \mid j = 1, 2, 3\}), \\ p^*(i, j) &= P(m_i^* + \mathbf{X}_i > m_j^* + \mathbf{X}_j), \\ p^*(i) &= P(m_i^* + \mathbf{X}_i = \max\{m_j^* + \mathbf{X}_j \mid j = 1, 2, 3\}), \end{aligned}$$

where the \mathbf{X}_i are independent identically distributed random variables with common distribution function $F(x) = e^{-e^{-x}}$ (Fig. 1 shows the corresponding probability density function). Moreover (i) implies that the ranking probabilities also satisfy the same double exponential GT-V model:

$$\begin{aligned} r(i, j, k) &= P(m_i + \mathbf{X}_i > m_j + \mathbf{X}_j > m_k + \mathbf{X}_k), \\ r^*(i, j, k) &= P(m_i^* + \mathbf{X}_i > m_j^* + \mathbf{X}_j > m_k^* + \mathbf{X}_k) \end{aligned}$$

(since $r(i, j, k) = p(j, k) - p(j)$ and $r^*(i, j, k) = p^*(j, k) - p^*(j)$.)

Now from the first part of the reversibility assumption we have

$$p(i, j) = P(\mathbf{X}_j - \mathbf{X}_i < m_i - m_j) = p^*(j, i) = P(\mathbf{X}_i - \mathbf{X}_j < m_j^* - m_i^*)$$

and consequently $m_i - m_j = m_j^* - m_i^*$. Since the m values are only unique up to addition of a constant we can set $m_1 = m_1^* = 0$, so that $m_i^* = -m_i$. Then the second part of reversibility holds iff

$$\begin{aligned} r(i, j, k) &= P(m_i + \mathbf{X}_i > m_j + \mathbf{X}_j > m_k + \mathbf{X}_k) \\ &= r^*(k, j, i) = p(-m_k + \mathbf{X}_k > -m_j + \mathbf{X}_j > -m_i + \mathbf{X}_i) \\ &= P(m_i - \mathbf{X}_i > m_j - \mathbf{X}_j > m_k - \mathbf{X}_k). \end{aligned}$$

This last equation is the key to the incompatibility between the Choice Axiom and reversibility. For suppose it were true for all possible values of m_1, m_2, m_3 —which is what would be required in order for the Choice Axiom to be generally consistent with reversibility. Then it would follow that the double exponential GT-V model would be equivalent for 3-alternative ranking experiments to the GT-V model based on the distribu-

tion function $F^*(x) = 1 - e^{-e^x}$, i.e., the distribution function of $-\mathbf{X}$ when \mathbf{X} is double exponential. However, in view of the result on ranking equivalence described earlier in Section 1.1, this could be true only if these two GT-V models were also equivalent for 3-alternative choice experiments, and this cannot be so because we already know that the double exponential GT-V model is unique for such experiments—i.e., the only GT-V models equivalent to the model based on $F(x) = e^{-e^{-x}}$ are those based on distributions of the same type, i.e., distribution functions of the form $e^{-e^{-(ax+b)}}$, where $a > 0$, b are arbitrary constants (Yellott, 1977). Moreover, because the models based on F and F^* would have to generate the same ranking systems using the same scale values (that is, $\{m_i + \mathbf{X}_i\}$ and $\{m_i - \mathbf{X}_i\}$ would have to generate the same ranking probabilities) the constant a here would have to be 1. (Footnote 3 below explains why.) Consequently reversibility could hold only if for some constant b

$$F^*(x) = 1 - F(-x) = 1 - e^{-e^{-x}} = F(x + b) = e^{-e^{-(x+b)}}$$

However, this would mean that the double exponential distribution is symmetric, i.e., for the centering constant $b/2$ (which would have to be zero here) the distribution functions $F(x + b/2)$ (corresponding to $\mathbf{X} - b/2$) and $1 - F(-x + b/2)$ (corresponding to $-(\mathbf{X} - b/2)$) would be identical. In other words, for the Choice Axiom to satisfy reversibility the densities corresponding to $1 - e^{-e^{-x}}$ and $e^{-e^{-x}}$ would have to be the same, and clearly they are not, as one can see in Fig. 1. However, one can also see from the figure that the asymmetry of the double exponential density is not very great (compared, for example, to an exponential distribution), and so it becomes understandable that the irreversibility implied by the Choice Axiom is numerically not very large.

Analyzed from the GT standpoint then, Block and Marschak's impossibility theorem appears not as an algebraic fluke, but rather as a natural consequence of the shape of the double exponential distribution—in particular, its asymmetry. However, the same analysis

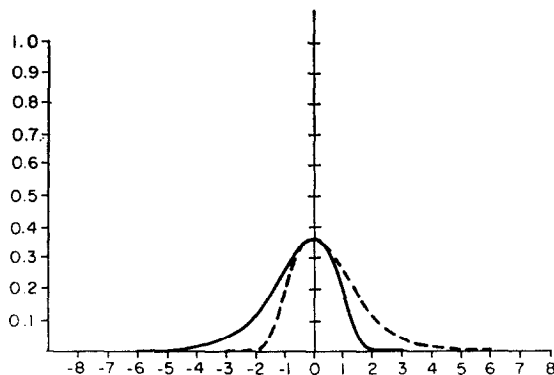


FIG. 1. Probability density functions for the double exponential distribution $F(x) = e^{-e^{-x}}$ (dashed line) and the "reverse" double exponential distribution $F^*(x) = 1 - e^{-e^{-x}}$ (solid line). The corresponding GT-V models are not equivalent for three alternative ranking experiments, and consequently neither is reversible for three (or more) alternatives.

also raises a general question. Suppose we retain the assumption that choices and rankings satisfy a common random utility model (i.e., (i)), but drop the decomposition assumption (ii)—which in this context was equivalent to assuming that both the best-to-worst and worst-to-best probabilities satisfy the double exponential GT-V model. Instead, suppose we replace (ii) with some other assumption that implies that both sets of probabilities satisfy some other GT-V model, i.e., a model based on some distribution other than the double exponential—normal, gamma, or whatever. Then the question arises, which of the resulting ranking models would satisfy the reversibility assumption (iii)?

This general problem is solved here in Section 3.4, using the ranking equivalence theorem of Section 3.3 and earlier results on choice equivalence summarized below in Section 2. On the basis of the special case of the double exponential model, one might expect that the answer would depend only on the symmetry of the utility distributions, i.e., one might expect that a GT-V model is reversible iff its utility distribution is symmetrical. However, this hunch turns out to be only half correct: Every symmetrical distribution does yield a reversible GT-V model, but so also do some asymmetric distributions—though in every case these curious models are only reversible for some finite number of alternatives and never for arbitrarily many. (For example, the asymmetric density function shown in the lower panel of Fig. 2 yields a GT-V model that is reversible for three alternatives but not for four.) The key to reversibility is indeed the shape of the utility distribution, but more is involved than symmetry alone: In addition, one needs to consider the shape of its Fourier transform.

1.3. Organization of the Paper

To put these results in context it seems useful to begin by reviewing the equivalence properties of GT models for choice experiments, since results on that problem that we need here are presently scattered through a series of papers. This is done in Section 2, and then Section 3 deals with ranking: 3.1 provides notation and definitions; 3.2 discusses the relationship between choice and ranking for random utility models in general; 3.3 gives the equivalence theorem for GT models for ranking; and 3.4 contains the results on reversibility.

2. REVIEW: EQUIVALENCE PROPERTIES OF GENERALIZED THURSTONE MODELS FOR CHOICE

2.1. Notation and Terminology

$A_n = \{a_1, a_2, \dots, a_n\}$ is a set of n choice alternatives, which we identify with the set of indices $I_n = \{1, 2, \dots, n\}$. In a *pair comparison* experiment alternatives are presented two at a time with the instruction “pick one”: Here $p(i, j)$ denotes the probability that a_i is chosen when $\{a_i, a_j\}$ is presented, and an entire collection of the form $\{p(i, j) \mid i, j \in I_n\}$ is called a *pair comparison system* for n alternatives. In a *complete choice experiment* one presents not only pairs, but all of the subsets of A_n : Here $p(i; S)$, $i \in S \subseteq I_n$, denotes the probability that a_i is chosen when the subject is presented with the subset $\{a_j \mid j \in S\}$ (and thus $p(i, j)$ is shorthand for $p(i, \{i, j\})$). A complete collection of probability distribu-

tions for all of the subsets of A_n (i.e., a collection of the form $\{\{p(i; S) \mid i \in S\} \mid S \subseteq I_n, |S| \geq 2\}$) is called a *complete system of choice probabilities* for n alternatives.

Two choice models (i.e., “theories”) are said to be *equivalent for pair comparison experiments* with n alternatives iff every pair comparison system for n alternatives that satisfies (either) one also satisfies the other. Similarly, two models are said to be *completely equivalent* for n alternatives iff every complete system of choice probabilities for n alternatives that satisfies one also satisfies the other.

2.2. Generalized Case V Models

Thurstone’s basic idea was, of course, to model choice behavior on the assumption that the alternatives a_1, a_2, \dots , correspond to real valued utility random variables U_1, U_2, \dots : When a subject is required to choose a single alternative from the subset $\{a_j \mid j \in S\}$ he picks the one with the largest utility random variable, so that $p(i; S) = P[U_i = \max\{U_j \mid j \in S\}]$. (Actually Thurstone only dealt explicitly with pair comparisons, but the extension to complete experiments is immediate.) For concreteness sake Thurstone assumed the U_i were normal random variables, and considered five special cases corresponding to various constraints on their variances and covariances. The simplest and most widely applied case, Case V, amounts to assuming that the U_i are independent and identically distributed (i.i.d.) except for shifts along the abscissa, i.e., they can be written in the form $U_1 = m_1 + X_1, U_2 = m_2 + X_2, \dots$, where the m_i are real constants (“scale values”) and the X_i are i.i.d. normal random variables.

Yellott (1977) considered a class of *Generalized Case V Thurstone (GT-V) Models* (called simply “Thurstone Models” in that paper) in which the X_i are i.i.d. with common distribution function F (i.e., $F(x) = P[X_i \leq x]$), subject to the condition that the difference distribution D_F (i.e., the distribution function of $X_i - X_j$) is everywhere continuous and strictly increasing. (This condition guarantees that $p(i, j)$ is a continuous strictly increasing function of the scale value difference $m_i - m_j$.) The GT-V model \mathcal{F}_F is then identified with the set of all systems of choice probabilities that can be generated by F together with arbitrary scale values m_1, m_2, \dots ; and the identifiability problem is whether \mathcal{F}_F and \mathcal{F}_G can be equivalent for some class of experiments even though F and G are different distributions. (Equivalence here means explicitly that any system of choice probabilities that can be generated by \mathcal{F}_F by assigning arbitrary values to the scale parameters m_1, m_2, \dots , can also be generated by \mathcal{F}_G using some set of parameters m'_1, m'_2, \dots ; and vice versa.) It is quickly apparent that if F and G are distributions of the same type (i.e., $F(x) = G(ax + b), a > 0$) \mathcal{F}_F and \mathcal{F}_G are equivalent for all choice experiments—in other words, choice experiments cannot identify the “true” mean and variance of F , but only (at most) its type: normal, exponential, gamma, or whatever. Consequently the interesting question is whether \mathcal{F}_F and \mathcal{F}_G can be equivalent when F and G are distributions of different types. The answer turns out to depend on the class of experiments one has in mind—in particular, on the number of alternatives.

2.2.1. *Pair comparisons.* Yellott (1977) shows that \mathcal{F}_F and \mathcal{F}_G are equivalent for pair comparison experiments with three or more alternatives (the smallest nontrivial choice experiment) iff they are equivalent for exactly three, and the latter is true iff $D_F(x) =$

$D_G(ax)$ for some $a > 0$.³ This condition is not very restrictive: Besides solutions of the form $F(x) = G(ax + b)$, and $F(x) = 1 - G(-ax + b)$ (where F is the distribution of \mathbf{X} and G the distribution of $-\mathbf{aX} + b$, $a > 0$ —so that F and G may or may not be of the same type, depending on whether they are symmetrical or asymmetrical), there are usually others in which F and G are not related in any probabilistically sensible way, but instead only through the Fourier analytic fact that their densities have essentially the same amplitude spectrum, i.e., the characteristic functions of F and G (denoted f and g , respectively) satisfy the relationship $|g(t)| = |f(at)|$. (In this respect pair comparison experiments are analogous to X-ray diffraction experiments in crystallography: Both can only identify amplitude spectra and necessarily lose the phase spectrum.) For example, the double exponential model $F(x) = e^{-e^{-x}}$ (which yields Luce's Choice Axiom) has pair comparison equivalents that are not of the same type as either $F(ax + b)$ or $1 - F(-ax + b)$, and this is also true of the exponential model $F(x) = 1 - e^{-x}$. The only exceptional case (reported so far) is the normal model (Thurstone's original Case V) whose pair comparison predictions cannot be entirely duplicated by any non-normal model—a special status that seems quite fitting, though historically entirely coincidental (Yellott, 1977, p. 131).

2.2.2. *Complete choice experiments.* The technical key here is the fact that \mathcal{F}_F and \mathcal{F}_G are equivalent for complete experiments with n alternatives iff the characteristic functions of F and G satisfy the relationship

$$g(t_1)g(t_2)\cdots g(t_{n-1})g\left(-\sum_{i=1}^{n-1}t_i\right) = f(at_1)f(at_2)\cdots f(at_{n-1})f\left(-a\sum_{i=1}^{n-1}t_i\right) \quad (1)$$

for some $a > 0$ and all t_1, t_2, \dots, t_{n-1} .⁴ (Yellott (1977) proves "only if"; Rockwell *et al.* (1978) supply the proof of "if," inadvertently omitted earlier.) $g(t) = e^{ibf(at)}$ is always a solution to (1); it corresponds to $F(x) = G(ax + b)$. And if (1) holds for all n , this is the only solution, i.e., \mathcal{F}_F and \mathcal{F}_G are completely equivalent for an infinite number of alternatives iff F and G are distributions of the same type.

Consequently if F and G are distributions of different types there is always some minimal number of alternatives necessary to distinguish between them. If F or G has a nonvanishing characteristic function, this number is three, i.e., if f or g is nonvanishing then for $n = 3$ the only solution to (1) is $g(t) = e^{ibf(at)}$ (Yellott, 1977). This condition is satisfied by all the well known distributions that are likely to occur to one as natural bases for GT-V models, e.g., the normal, gamma, double exponential, etc., since all have non-

³ The constant a determines the relationship between the scale values of \mathcal{F}_F and \mathcal{F}_G : If \mathcal{F}_F generates a given pair comparison system with scale values m_1, m_2, \dots, m_n ; then the corresponding scale values for \mathcal{F}_G , say m'_1, m'_2, \dots, m'_n , must satisfy the condition $m'_i - m'_j = a(m_i - m_j)$. (Since $p(i, j) = D_F(m_i - m_j) = D_G(a(m_i - m_j)) = D_G(m'_i - m'_j)$.) Note that if there exists a system of pair comparison probabilities that is generated by both D_F and D_G with the same set of scale values (i.e., some set m_1, \dots, m_n yields the same system under both models) and $m_i \neq m_j$ for at least one pair $i, j, i \neq j$, then a must be 1 (since then $D_F(m_i - m_j) = D_G(a(m_i - m_j)) = D_G(m_i - m_j)$).

⁴ a here has the same interpretation as in Footnote 3.

vanishing characteristic functions (it's worth recalling that this is true of all infinitely divisible distributions). Consequently one can say that all of these models are in principle identifiable by the smallest possible complete choice experiments.

However, nothing in their definition forces GT-V models to have nonvanishing characteristic functions, and some do not. In those cases (1) can have solutions f and g that correspond to distributions of different types—for $n = 3$ (Yellott, 1978; Moszner, 1978), and indeed for any n (Rockwell *et al.*, 1979). Consequently for every n one can construct pairs of GT-V models that are completely equivalent for experiments with n (or fewer) alternatives but that become nonequivalent for $n + 1$. (Specific examples are the models corresponding to the density functions $\text{sinc}^2(x)(1 + \cos 2\pi nx)$ and $\text{sinc}^2(x)(1 + \sin 2\pi nx)$, where $\text{sinc}(x) = (\sin \pi x)/\pi x$. Figure 2 illustrates these densities for $n = 3$.) These cases prompted the present analysis of ranking experiments, since it seemed possible that the extra information provided by ranking probabilities might enable one to distinguished between models on the basis of fewer alternatives.

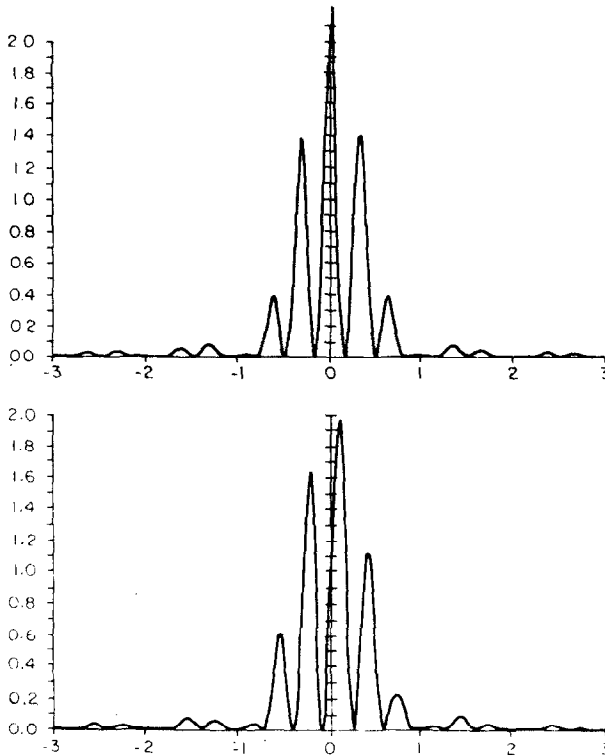


FIG. 2. Top panel shows the probability density function $[\text{sinc}^2(x) (1 + \cos 6\pi x)]$. Bottom panel shows the density $[\text{sinc}^2(x) (1 + \sin 6\pi x)]$. The corresponding GT-V models are equivalent for both choice and ranking experiments with three alternatives, and both models are reversible for three alternative ranking experiments despite the asymmetry of the lower density (Corollary 3, Sect. 3.4).

2.3. Generalized Thurstone (GT) Models

Strauss (1979) deals with this broader class of models, which includes the GT-V models as a special case. A GT model \mathcal{F}_n corresponds to specific sequence of n random variables $\mathbf{X}_1, \dots, \mathbf{X}_n$ with joint distribution function $F_n(x_1, x_2, \dots, x_n) = P(\mathbf{X}_1 \leq x_1, \mathbf{X}_2 \leq x_2, \dots, \mathbf{X}_n \leq x_n)$. For the same reason that the difference distribution D_F is assumed to be continuous and strictly increasing in the case of GT-V models, F_n here is assumed to satisfy the condition that for all i and j , $i \neq j$, the difference distribution $P(\mathbf{X}_i - \mathbf{X}_j \leq x)$ is everywhere continuous and strictly increasing. Otherwise F_n is unrestricted—in particular, nothing is assumed about dependencies between the \mathbf{X}_i . A complete system of choice probabilities for n alternatives is said to satisfy \mathcal{F}_n iff there exist scale values m_1, m_2, \dots, m_n such that $\forall i \in S \subseteq I_n, p(i; S) = P(m_i + \mathbf{X}_i = \max\{m_j + \mathbf{X}_j \mid j \in S\})$. (The model \mathcal{F}_n is not defined for experiments with more than n alternatives. Consequently each of the GT-V models of Section 2.2 corresponds in this context to an infinite sequence of GT models of the form $\mathcal{F} = \{\mathcal{F}_n \mid n = 2, 3, \dots\}$, where $F_n(x_1, \dots, x_n) = \prod_{i=1}^n F(x_i)$. Two GT models \mathcal{F}_n and \mathcal{G}_n are completely equivalent iff every complete system that satisfies one also satisfies the other.)

Naturally in this broader class of models the equivalence possibilities are greatly enriched. Nevertheless Strauss was able to show that identifiability here depends on a straightforward generalization of the Fourier analytic criterion for GT-V models expressed in Eq. 1:

LEMMA 1. *Suppose $f_n(t_1, \dots, t_n)$ and $g_n(t_1, \dots, t_n)$ are the joint characteristic functions of F_n and G_n , respectively. Then the GT models \mathcal{F}_n and \mathcal{G}_n are completely equivalent iff*

$$g_n(t_1, \dots, t_n) = f_n(at_1, \dots, at_n) \quad (2)$$

for some $a > 0$ and all t_1, \dots, t_n such that $\sum_{i=1}^n t_i = 0$.

Proof (Strauss, 1979, Theorem 1). Note that (2) reduces to (1) in the case of GT-V models.

This result also proves to be the key to the equivalence properties of GT models for ranking, as shown below in Section 3.3.

3. GENERALIZED THURSTONE MODELS FOR RANKING

3.1. Terminology and Notation

As before, A_n denotes a set of alternatives $\{a_1, \dots, a_n\}$; I_n the corresponding set of indices $\{1, \dots, n\}$. In a ranking experiment we present subsets of A_n and ask the subject to rank order the alternatives in each subset from best to worst: The best alternative is assigned rank 1, the next best rank 2, and so on. The result for each subset is a probability distribution over its possible rank orderings, i.e., over the possible permutations of its indices. To denote this, suppose $S \subseteq I_n$ and ρ is a permutation of the indices in S , i.e., ρ maps S onto the integers $1, 2, \dots, |S|$: $\rho(j)$ is the ordinal position assigned to alternative

j , and $\rho^{-1}(i)$ is the index of the alternative that has rank order i . Then $r(\rho; S)$ denotes the probability of ranking $\{a_j \mid j \in S\}$ in the order $\langle a_{\rho^{-1}(1)}, a_{\rho^{-1}(2)}, \dots, a_{\rho^{-1}(|S|)} \rangle$. To indicate the ranking explicitly, let $i_\rho = \rho^{-1}(i)$: Then we sometimes write $r(1_\rho, 2_\rho, \dots, \mid S \mid_\rho)$ for $r(\rho; S)$. For example, $r(4, 1, 2)$ is the probability of ranking the alternative set $\{a_1, a_2, a_4\}$ in the order $\langle a_4, a_1, a_2 \rangle$.

An experiment that determines $r(\rho; S)$ for every rank ordering of every subset of I_n is a complete ranking experiment, and the resulting collection of probability distributions is a complete system of ranking probabilities for n alternatives. Such a system satisfies the GT model \mathcal{F}_{F_n} iff there exist scale values m_1, \dots, m_n such that for every ranking ρ of every $S, S \subseteq I, \mid S \mid \geq 2$:

$$r(1_\rho, 2_\rho, \dots, \mid S \mid_\rho) = P[m_{1_\rho} + \mathbf{X}_{1_\rho} > m_{2_\rho} + \mathbf{X}_{2_\rho} > \dots > m_{\mid S \mid_\rho} + \mathbf{X}_{\mid S \mid_\rho}].$$

In other words a system of ranking probabilities satisfies \mathcal{F}_{F_n} iff it can be generated by assigning utility random variables $m_1 + \mathbf{X}_1, \dots, m_n + \mathbf{X}_n$ to a_1, \dots, a_n (with F_n the joint distribution function of $\mathbf{X}_1, \dots, \mathbf{X}_n$) and applying the decision rule: "For every subset $\{a_j \mid j \in S\}$ rank the alternatives in the same order as the random variables $\{m_j + \mathbf{X}_j \mid j \in S\}$." Then we say that two GT models are equivalent for ranking iff every complete system of ranking probabilities that satisfies one also satisfies the other, and the identifiability problem is to determine when this can happen for two models that correspond to nontrivially different distributions (i.e., it's obvious that \mathcal{F}_{F_n} and \mathcal{F}_{G_n} are equivalent for ranking as well as for choice if

$$F_n(x_1, \dots, x_n) = G_n(ax_1 + b, \dots, ax_n + b),$$

where $a > 0, b$ are arbitrary constants, and so the question is whether there are other possibilities.)

In particular, since it is immediately clear that for GT models ranking equivalence implies choice equivalence, it is natural to wonder whether two such models can be equivalent for choice but not for ranking. To put this question in perspective it is useful to begin by considering relationships between choice and ranking for the broadest class of models that still retain Thurstone's basic idea: These are the "random utility models" described in the next section.⁵ This is important because the assumption that choices and rankings are both dictated by order relationships within a common set of random variables obviously implies a structural bond between the two systems of preference probabilities, and so it is necessary to understand this before considering the additional implications of the distributional constraints embodied in GT models.

3.2. Random utility models for choice and ranking: Constraints on equivalence relationships.

By a random utility vector \mathbf{U}_n we mean simply a sequence of jointly distributed random variables $(\mathbf{U}_1, \dots, \mathbf{U}_n)$, subject to the constraint $P(\mathbf{U}_i = \mathbf{U}_j) = 0$ for all $i \neq j$. Such a

⁵ Luce and Suppes (1965); Corbin and Marley (1974); Falmagne (1978), Cohen (1980).

vector *generates* a complete system of choice probabilities $\{p(i; S) \mid S \subseteq I_n\}$ via the rule

$$p(i; S) = P[\mathbf{U}_i = \text{Max}\{\mathbf{U}_j \mid j \in S\}]$$

and a complete system of ranking probabilities $\{r(i; S) \mid S \subseteq I_n\}$ via the rule

$$r(i; S) = P[\mathbf{U}_{1_p} > \mathbf{U}_{2_p} > \cdots > \mathbf{U}_{|S|_p}].$$

By a *random utility (RU) model* \mathcal{U} we mean any collection of random utility vectors: A given system of choice or ranking probabilities is said to satisfy \mathcal{U} iff it is generated by some vector in \mathcal{U} . Two RU models \mathcal{U} and \mathcal{U}' are equivalent for choice (ranking) experiments with n alternatives iff every complete system of choice (ranking) probabilities for n alternatives that satisfies one also satisfies the other.

Clearly every GT model is a RU model, and likewise every GT-V model (construed as a sequence of GT models, as in Section 2.3). However, a RU model may also consist of just one random utility vector.

Now suppose \mathcal{U} and \mathcal{U}' are equivalent for ranking experiments with n alternatives, and that some system of choice probabilities $\{p(i; S) \mid S \subseteq I_n\}$ satisfies \mathcal{U} . Then this system is generated by some vector $\bar{\mathbf{U}}_n$ in \mathcal{U} , and that vector simultaneously generates a system of ranking probabilities $\{r(i; S) \mid S \subseteq I_n\}$ which also must satisfy \mathcal{U}' for some vector $\bar{\mathbf{U}}'_n$. This vector in turn must generate the same choice system as $\bar{\mathbf{U}}_n$, since the choice probabilities generated by a given utility vector are completely determined by its ranking probabilities via the relationship

$$p(i; S) = \sum_{\rho: 1_\rho=i} r(\rho; S),$$

i.e., the probability of choosing i from S is the probability of ranking i first in S . Consequently

LEMMA 2. *If two RU models are equivalent for ranking experiments with n alternatives they are equivalent for choice experiments with n alternatives.*

Next, suppose \mathcal{U} and \mathcal{U}' are equivalent for choice experiments with *three* alternatives, and $\{r(i; S) \mid S \subseteq I_3\}$ is a ranking system generated by $\bar{\mathbf{U}}_3 \in \mathcal{U}$. Then this vector generates a choice system $\{p(i; S) \mid S \subseteq I_3\}$, and that system must also be generated by some vector in \mathcal{U}' , say $\bar{\mathbf{U}}'_3$. That vector in turn must generate the same ranking system as $\bar{\mathbf{U}}_3$, because for three alternatives the ranking probabilities generated by a utility vector are completely determined by the corresponding choice probabilities, via the relationships

$$r(i, j) = P[\mathbf{U}_i > \mathbf{U}_j] = p(i, j),$$

$$\begin{aligned} r(i, j, k) &= P[\mathbf{U}_i > \mathbf{U}_j > \mathbf{U}_k] = P[\mathbf{U}_j > \mathbf{U}_k] - P[\mathbf{U}_j = \text{Max}\{\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3\}] \\ &= p(j, k) - p(j; \{1, 2, 3\}). \end{aligned}$$

Combining this argument with Lemma 2, we see that:

LEMMA 3. *Two RU models are equivalent for choice experiments with three alternatives iff they are equivalent for ranking experiments with three alternatives.*

For experiments with four or more alternatives, however, the ranking probabilities generated by a random utility vector are no longer determined by its choice probabilities. Table 1 illustrates two different probability distributions over the possible rank ordering of four alternatives that yield identical systems of choice probabilities. Each of these ranking probability distributions is generated by a random utility vector, namely the vector (U_1, U_2, U_3, U_4) having the joint distribution

$$P[U_1 = i^{-1}, U_2 = j^{-1}, U_3 = k^{-1}, U_4 = l^{-1}] = P[\rho(1) = i, \rho(2) = j, \rho(3) = k, \rho(4) = l],$$

TABLE 1
Random Utility Models That Are Equivalent
for Choice but Not for Ranking

Rank order	P[ranking] for Model 1	P[ranking] for Model 2
1234	a^*	b
1243	b^*	a
2134	b	a
2143	a	b
1324	a	b
1342	b	a
3124	b	a
3142	a	b
1423	a	b
1432	b	a
4123	b	a
4132	a	b
2314	a	b
2341	b	a
3214	b	a
3241	a	b
2413	a	b
2431	b	a
4213	b	a
4231	a	b
4321	a	b
4312	b	a
3421	b	a
3412	a	b

* $a + b = 1/12, a \neq b.$

where i, j, k, l is any permutation of 1, 2, 3, 4. Since the left and right columns in Table 1 assign different probabilities to every rank order they obviously correspond to different utility vectors, say \mathbf{U}_3 and \mathbf{U}'_3 , but these vectors both generate the same system of choice probabilities (i.e., $p(i; S) \equiv 1/|S|$). Consequently the two systems of rank order probabilities in Table 1 correspond to two random utility models $\{\mathbf{U}_3\}$ and $\{\mathbf{U}'_3\}$ that are not equivalent for ranking experiments with four alternatives but are equivalent for choice experiments with four alternatives.

For RU models in general then, ranking equivalence implies choice equivalence, but not conversely—except in the special case of experiments with three alternatives. Therefore it is reasonable to ask whether there are GT models that are equivalent for complete choice experiments with $n \geq 4$ alternatives but not for the corresponding ranking experiments. The next section shows that there are none.

3.3. Equivalence Properties of GT Models for Ranking

THEOREM 1. *Two Generalized Thurstone Models are equivalent for ranking experiments with n alternatives iff they are equivalent for choice experiments with n alternatives.*

Proof. “Only if” follows immediately from Lemma 2, since every GT model is a RU model.

To show “if,” suppose \mathcal{F}_n and \mathcal{G}_n are equivalent for choice experiments and that $\{r(\cdot; S) \mid S \subseteq I_n\}$ is a complete system of ranking probabilities that satisfies \mathcal{F}_n for some set of scale values m_1, m_2, \dots, m_n . From Strauss’ theorem (Lemma 1) we know that for some $a > 0$ the characteristic functions of F_n and G_n satisfy

$$f_n(t_1, t_2, \dots, t_n) = g_n(t_1/a, t_2/a, \dots, t_n/a) \quad (3)$$

whenever $\sum_{i=1}^n t_i = 0$. We will show that \mathcal{G}_n generates $\{r(\cdot; S) \mid S \subseteq I_n\}$ with scale values am_1, am_2, \dots, am_n . Note first that we need only show that \mathcal{G}_n can generate the same ranking probabilities for the entire set A_n (i.e., the same values of $r(\rho; I_n)$), since these completely determine the ranking probabilities for all the proper subsets of A_n . Now let $m_1 + \mathbf{X}_1, \dots, m_n + \mathbf{X}_n$ be the utility random variables of \mathcal{F}_n and $am_1 + \mathbf{X}'_1, \dots, am_n + \mathbf{X}'_n$ those of \mathcal{G}_n , and suppose that ρ is any permutation of 1, 2, ..., n . Then we need to show that

$$\begin{aligned} P[m_{1_\rho} + \mathbf{X}_{1_\rho} > m_{2_\rho} + \mathbf{X}_{2_\rho} > \dots > m_{n_\rho} + \mathbf{X}_{n_\rho}] \\ = P[am_{1_\rho} + \mathbf{X}'_{1_\rho} > am_{2_\rho} + \mathbf{X}'_{2_\rho} > \dots > am_{n_\rho} + \mathbf{X}'_{n_\rho}]. \end{aligned} \quad (4)$$

Now the left side here is

$$\begin{aligned} P[\mathbf{X}_{2_\rho} - \mathbf{X}_{1_\rho} < m_{1_\rho} - m_{2_\rho}, \\ \mathbf{X}_{3_\rho} - \mathbf{X}_{2_\rho} < m_{2_\rho} - m_{3_\rho}, \dots, \mathbf{X}_{n_\rho} - \mathbf{X}_{n-1_\rho} < m_{n-1_\rho} - m_{n_\rho}] \end{aligned}$$

and the right side is

$$P[\mathbf{X}'_{2_\rho} - \mathbf{X}'_{1_\rho} < a(m_{1_\rho} - m_{2_\rho}), \\ \mathbf{X}'_{3_\rho} - \mathbf{X}'_{2_\rho} < a(m_{2_\rho} - m_{3_\rho}), \dots, \mathbf{X}'_{n_\rho} - \mathbf{X}'_{n-1_\rho} < a(m_{n-1_\rho} - m_{n_\rho})]$$

and so (4) is true if the random vectors

$$(\mathbf{X}_{2_\rho} - \mathbf{X}_{1_\rho}, \mathbf{X}_{3_\rho} - \mathbf{X}_{2_\rho}, \dots, \mathbf{X}_{n_\rho} - \mathbf{X}_{n-1_\rho}) \quad (5)$$

and

$$((\mathbf{X}'_{2_\rho} - \mathbf{X}'_{1_\rho})/a, (\mathbf{X}'_{3_\rho} - \mathbf{X}'_{2_\rho})/a, \dots, (\mathbf{X}'_{n_\rho} - \mathbf{X}'_{n-1_\rho})/a) \quad (6)$$

have the same joint distributions. The joint characteristic function of (5) is

$$E[\exp(is_1(\mathbf{X}_{2_\rho} - \mathbf{X}_{1_\rho}) + is_2(\mathbf{X}_{3_\rho} - \mathbf{X}_{2_\rho}) + \dots + is_{n-1}(\mathbf{X}_{n_\rho} - \mathbf{X}_{n-1_\rho}))] \\ = E[\exp(-is_1\mathbf{X}_{1_\rho} + i(s_1 - s_2)\mathbf{X}_{2_\rho} + \dots + i(s_{n-2} - s_{n-1})\mathbf{X}_{n-1_\rho} + is_{n-1}\mathbf{X}_{n_\rho})]$$

and if we write $t_1 = -s_1$, $t_2 = s_1 - s_2$, ..., $t_{n-1} = s_{n-2} - s_{n-1}$, $t_n = s_{n-1}$, this last expression becomes

$$f_n(t_{\rho(1)}, t_{\rho(2)}, \dots, t_{\rho(n)}), \quad (7)$$

where $\sum_{i=1}^n t_{\rho(i)} = 0$. Similarly the joint characteristic function of (6)

$$E[\exp(is_1(\mathbf{X}'_{2_\rho} - \mathbf{X}'_{1_\rho})/a + \dots + is_{n-1}(\mathbf{X}'_{n_\rho} - \mathbf{X}'_{n-1_\rho})/a)]$$

becomes

$$g_n(t_{\rho(1)}/a, t_{\rho(2)}/a, \dots, t_{\rho(n)}/a) \quad (8)$$

and (7) and (8) must be equal because of (3). Thus the random vectors (5) and (6) have the same characteristic function, and consequently are identically distributed. This establishes (4) and completes the proof. ■

In particular we see that two GT-V models are completely equivalent for ranking experiments with n alternatives iff they are completely equivalent for choice experiments with n alternatives, and so, for example, the models corresponding to the density functions $\text{Sinc}^2(x)(1 + \text{Cos } 2\pi nx)$ and $\text{Sinc}^2(x)(1 + \text{Sin } 2\pi nx)$ cannot be discriminated by either choice or ranking experiments with $n' \leq n$ alternatives.

3.4. Reversibility

In the Introduction (Section 1.2) the concept of reversibility was motivated in terms of Block and Marschak's impossibility theorem. It was pointed out that this theorem impli-

citly assumes that both the best-to-worst and worst-to-best ranking probabilities satisfy the double exponential GT-V model, and consequently reversibility could hold generally only if that model were equivalent (for three-alternative ranking experiments) to the GT-V model based on the “reverse” double exponential distribution—i.e., the distribution of $-\mathbf{X}$ when \mathbf{X} itself is a double exponential random variable. Because of Theorem 1 in the last section we know that this ranking equivalence could only hold if these two models were also equivalent for three-alternative choice experiments, and consequently we can interpret the impossibility theorem in terms of the general equivalence properties of GT-V choice models, as outlined in Section 2.2. In particular, because the double exponential distribution has a nonvanishing characteristic function we know that a GT-V model \mathcal{T}_G can be equivalent to the model \mathcal{T}_F based on the double exponential distribution $F(x) = e^{-e^{-x}}$ iff G is another distribution of the same type, i.e., $G(x) = F(ax + b)$ for some $a > 0$. Since the double exponential distribution is asymmetric, the reverse double exponential random variable $-\mathbf{X}$ with distribution function $F^*(x) = 1 - e^{-e^x}$ cannot be of double exponential type, and so \mathcal{T}_{F^*} cannot be equivalent to \mathcal{T}_F for three alternative choice experiments. Consequently one can say that the impossibility theorem is ultimately due to the shape of the double exponential distribution, i.e., its asymmetry and its “smoothness,” as reflected in the fact that its characteristic function is nonvanishing.

This analysis then suggests the general problem: which GT models are reversible for n -alternative ranking experiments? To answer this question we first need a general definition of reversibility. Intuitively, reversibility means that every system of ranking probabilities $\{r(\rho); S \mid S \subseteq I_n\}$ generated by \mathcal{T}_{F_n} is matched by another system $\{r^*(\rho; S) \mid S \subseteq I_n\}$, also generated by \mathcal{T}_{F_n} , which assigns the same probabilities to the reverse rankings, that is, if ρ is any permutation of the indices in S and ρ^* is the reversed permutation (say $S = \{1, 2, 4\}$, $\rho(1) = 3$, $\rho(2) = 2$, $\rho(4) = 1$; then $\rho^*(1) = 1$, $\rho^*(2) = 2$, $\rho^*(4) = 3$), then $r^*(\rho^*) = r(\rho)$. To formalize this notion for GT models we will say that:

A GT model \mathcal{T}_{F_n} is reversible for n alternative ranking experiments iff for every set of scale values m_1, m_2, \dots, m_n there is another set $m_1^, m_2^*, \dots, m_n^*$ such that for every permutation ρ of every $S \subseteq I_n$*

$$\begin{aligned}
 &P[m_{1\rho} + \mathbf{X}_{1\rho} > m_{2\rho} + \mathbf{X}_{2\rho} > \dots > m_{|S|\rho} + \mathbf{X}_{|S|\rho}] \\
 &= P[m_{1S_{1\rho}}^* + \mathbf{X}_{|S|\rho} > m_{|S|-1\rho}^* + \mathbf{X}_{|S|-1\rho} > \dots > m_{1\rho}^* + \mathbf{X}_{1\rho}], \tag{9}
 \end{aligned}$$

where $\mathbf{X}_1, \dots, \mathbf{X}_n$ have the joint distribution function F_n .

Next we relate reversibility and equivalence:

LEMMA 4. *If F_n is the joint distribution function of the random variables $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$, let F_n^* denote the joint distribution function of $-\mathbf{X}_1, -\mathbf{X}_2, \dots, -\mathbf{X}_n$. Then the GT model \mathcal{T}_{F_n} is reversible for n alternative ranking experiments iff \mathcal{T}_{F_n} is equivalent to $\mathcal{T}_{F_n^*}$.*

Proof. Suppose \mathcal{T}_{F_n} and $\mathcal{T}_{F_n^*}$ are equivalent. Then for every system of ranking probabilities generated by \mathcal{T}_{F_n} (using some set of scale values m_1, m_2, \dots, m_n) there exists

a corresponding set $m_1^*, m_2^*, \dots, m_n^*$ by which $\mathcal{F}_{F_n^*}$ generates the same system, i.e., for every permutation ρ of every $S \subseteq I_n$

$$\begin{aligned} P[m_{1\rho} + \mathbf{X}_{1\rho} > \dots > m_{|S|\rho} + \mathbf{X}_{|S|\rho}] \\ &= P[m_{1\rho}^* - \mathbf{X}_{1\rho} > \dots > m_{|S|\rho}^* - \mathbf{X}_{|S|\rho}] \\ &= P[-m_{|S|\rho}^* + \mathbf{X}_{|S|\rho} > \dots > -m_{1\rho}^* + \mathbf{X}_{1\rho}] \end{aligned}$$

so \mathcal{F}_{F_n} is reversible. Conversely, suppose \mathcal{F}_{F_n} is reversible. Then any system of ranking probabilities generated by \mathcal{F}_{F_n} using some set of scale values m_1, \dots, m_n is also generated by $\mathcal{F}_{F_n^*}$ using scale values $-m_1^*, \dots, -m_n^*$, where the m_i^* are the scale values whose existence is guaranteed by the definition of reversibility, and so \mathcal{F}_{F_n} and $\mathcal{F}_{F_n^*}$ are equivalent. ■

Now we can give a general characterization of reversible GT models:

THEOREM 2. *A GT model \mathcal{F}_{F_n} is reversible iff the characteristic function f_n of distribution F_n satisfies the equation*

$$f_n(t_1, t_2, \dots, t_n) = f_n(-t_1, -t_2, \dots, -t_n) \tag{10}$$

for all t_i such that $\sum_{i=1}^n t_i = 0$.

Proof. Suppose (10) holds for all t_i such that $\sum_{i=1}^n t_i = 0$. Then since $f_n(t_1, \dots, t_n)$ is the characteristic function of F_n and $f_n(-t_1, \dots, -t_n)$ is the characteristic function of F_n^* , it follows from Strauss' theorem (Lemma 1, Section 2.3) that \mathcal{F}_{F_n} and $\mathcal{F}_{F_n^*}$ are equivalent for choice, and consequently (by Theorem 1) they are equivalent for ranking. Consequently (by Lemma 2) \mathcal{F}_{F_n} is reversible.

Conversely, suppose \mathcal{F}_{F_n} is reversible. Then from Lemma 2 we know that \mathcal{F}_{F_n} is equivalent to $\mathcal{F}_{F_n^*}$ for ranking, and consequently (Theorem 1) for choice. Therefore from Lemma 1 we have

$$f_n(t_1, t_2, \dots, t_n) = f_n(-at_1, \dots, -at_n) \tag{11}$$

for some $a > 0$ and all t_i such that $\sum t_i = 0$. To show that a here must be one (i.e., to show that (11) implies (10)) note that $f(-1, l, 0, 0, \dots, 0)$ is the characteristic function of $\mathbf{X}_2 - \mathbf{X}_1$, and $f(at, -at, 0, 0, \dots, 0)$ is the characteristic function of $a(\mathbf{X}_1 - \mathbf{X}_2)$. Consequently (11) implies that these random variables are identically distributed, i.e.,

$$\begin{aligned} P[\mathbf{X}_2 - \mathbf{X}_1 \leq y] &= P[\mathbf{X}_1 - \mathbf{X}_2 \leq y/a] \\ &= 1 - P[\mathbf{X}_2 - \mathbf{X}_1 \leq -y/a]. \end{aligned}$$

Let $D(y) = P[\mathbf{X}_2 - \mathbf{X}_1 \leq y]$, and recall that by the definition of GT models, D is strictly increasing. Now set $y = a$, to obtain $D(a) = 1 - D(-1)$, and then $y = -1$,

to obtain $D(-1) = 1 - D(1/a)$. Combining these two equations we have $D(a) = D(1/a)$, and since D is strictly increasing, a must be 1. ■

Specializing Theorem 2 to the case of GT-V models, where $f_n(t_1, t_2, \dots, t_n) = f(t_1)f(t_2) \cdots f(t_n)$, we have the immediate

COROLLARY 1. *A GT-V model \mathcal{F}_F is reversible for n alternatives iff the characteristic function f of distribution F satisfies the equation*

$$f(t_1)f(t_2) \cdots f(t_{n-1})f\left(-\sum_{i=1}^{n-1} t_i\right) = f(-t_1)f(-t_2) \cdots f(-t_{n-1})f\left(\sum_{i=1}^{n-1} t_i\right). \quad (12)$$

Proof. Equation (10) reduces to (12) in this case.

From this we can conclude that GT-V models with nonvanishing characteristic functions are only reversible in the obvious way, i.e., when their utility distributions are symmetric:

COROLLARY 2. *If F has a nonvanishing characteristic function the GT-V model \mathcal{F}_F is reversible for three or more alternatives iff F is a symmetric distribution (i.e., for some centering constant c , $F(x+c) = 1 - F(-x+c)$).*

Proof. Suppose F is symmetric; i.e., for some c , $\mathbf{X} - c$ and $-(\mathbf{X} - c)$ are identically distributed. The characteristic function of the former is $e^{-ict}f(t)$, and of the latter, $e^{ict}f(-t)$, and so these two functions are equal for all t , i.e., $f(t) = e^{i2ct}f(-t)$. Consequently for every n

$$\begin{aligned} f(t_1) \cdots f(t_{n-1})f\left(-\sum_{i=1}^{n-1} t_i\right) \\ &= e^{i2ct_1}f(-t_1) \cdots e^{i2ct_{n-1}}f(-t_{n-1}) e^{-i2c\sum_{i=1}^{n-1} t_i} f\left(\sum_{i=1}^{n-1} t_i\right) \\ &= f(-t_1) \cdots f(-t_{n-1})f\left(\sum_{i=1}^{n-1} t_i\right) \end{aligned}$$

so (12) holds and \mathcal{F}_F is reversible for any n .

Conversely if \mathcal{F}_F is reversible for $n = 3$ then (12) holds for $n = 3$ and so the characteristic functions $f(t)$ and $f^*(t) = f(-t)$ satisfy the equation

$$f(t_1)f(t_2)f(-t_1 - t_2) = f^*(t_1)f^*(t_2)f^*(-t_1 - t_2).$$

As noted earlier in Section 2.2, Yellott (1977) shows that when f is nonvanishing, this equation implies $f^*(t) = e^{ibt}f(t)$, so that here $f(-t) = e^{ibt}f(t)$. Consequently $-\mathbf{X}$ and $\mathbf{X} + b$ are identically distributed, and so for the centering constant $c = -b/2$ we have symmetry, i.e., $F(x + b/2) = 1 - F(-x + b/2)$. ■

Corollary 2 shows that for GT-V models with non-vanishing characteristic functions, reversibility and symmetry are equivalent conditions. However, if the characteristic

function of a distribution F is allowed to have zeros, \mathcal{F}_F can be reversible even though F is asymmetric. This can happen if \mathcal{F}_F is equivalent to a symmetric model, as in the following example.

COROLLARY 3. *The GT-V model corresponding to the asymmetric probability density function*

$$p(x) = \text{sinc}^2(x)(1 + \sin 2\pi nx)$$

is reversible for $n' \leq n$ alternatives, but not for $n' > n$.

Proof. The characteristic function $f(t)$ of $p(x)$ is

$$\begin{aligned} f(t) &= 1 - |t| && t \in (-1, 1) \\ &= -\frac{i}{2}(1 - |t - n|) && t \in (n - 1, n + 1) \\ &= \frac{i}{2}(1 - |t + n|) && t \in (-n - 1, -n + 1) \\ &= 0 && \text{elsewhere.} \end{aligned}$$

Rockwell *et al.* (1979) show that for $n' \leq n$, this characteristic function satisfies the equation

$$f(t_1)f(t_2) \cdots f(t_{n'-1})f\left(-\sum_{i=1}^{n'-1} t_i\right) = g(t_1)g(t_2) \cdots g(t_{n'-1})g\left(-\sum_{i=1}^{n'-1} t_i\right),$$

where g is the characteristic function of the symmetric density $q(x) = \text{sinc}^2(x)(1 + \cos 2\pi nx)$, given explicitly by

$$\begin{aligned} g(t) &= \begin{cases} 1 - |t| & t \in (-1, 1) \\ \frac{1}{2}(1 - |t - n|) & t \in (n - 1, n + 1) \\ \frac{1}{2}(1 - |t + n|) & t \in (-n - 1, -n + 1) \\ 0 & \text{elsewhere.} \end{cases} \end{aligned}$$

Now since density q is symmetric, it follows from the proof of Corollary 2 that for all n'

$$g(t_1) \cdots g(t_{n'-1})g\left(-\sum_{i=1}^{n'-1} t_i\right) = g(-t_1) \cdots g(-t_{n'-1})g\left(\sum_{i=1}^{n'-1} t_i\right)$$

and so for $n' \leq n$

$$\begin{aligned} f(-t_1) \cdots f(-t_{n'-1})f\left(\sum_{i=1}^{n'-1} t_i\right) &= g(-t_1) \cdots g(-t_{n'-1})g\left(\sum_{i=1}^{n'-1} t_i\right) \\ &= g(t_1) \cdots g(t_{n'-1})g\left(-\sum_{i=1}^{n'-1} t_i\right) \\ &= f(t_1) \cdots f(t_{n'-1})f\left(-\sum_{i=1}^{n'-1} t_i\right). \end{aligned}$$

Consequently f satisfies (12) for $n' \leq n$, and so the GT-V model based on density p is reversible for $n' \leq n$. However, for $n' \geq n + 1$ (12) fails, as one can see for $n' = n + 1$ by setting $t_1 = t_2 = \dots = t_{n'-1} = k$, where $(n - 1)/n < k < 1$. Then $f(1) = f(-1)$, since k falls in the interval $(-1, 1)$, but $-\sum_{i=1}^{n'-1} t_i = -nk$ falls in the interval $(-n - 1, -n + 1)$ while $\sum_{i=1}^{n'-1} t_i$ falls in the interval $(n - 1, n + 1)$. Consequently in this case $f(t_1) \dots f(t_{n-1}) = f(-t_1) \dots f(-t_{n-1}) > 0$, but

$$f\left(-\sum_{i=1}^{n'-1} t_i\right) \neq f\left(\sum_{i=1}^{n'-1} t_i\right)$$

since one side has the sign of $-i$, the other the sign of $+i$. A similar argument works for any $n' \geq n + 1$ and consequently (12) cannot hold in these cases, so the GT-V model based on density p is not reversible for more than n alternatives. ■

Finally, we note that although GT-V models based on asymmetric distributions can sometimes be reversible for a finite number of alternatives (as in the last corollary), they can never be reversible for arbitrarily many: There is always some number beyond which reversibility fails. Consequently if one wants to insist on reversibility for all possible ranking experiments, only symmetric models will do.

COROLLARY 4. *A GT-V model T_F is reversible for n alternatives for every n iff F is a symmetric distribution.*

Proof. Corollary 2 shows that if F is symmetric, \mathcal{T}_F is reversible for every n . Conversely if \mathcal{T}_F is reversible for every n then (12) holds for all n . Yellott (1977, Theorem 3) shows that this implies $f(t) = e^{ibt}f(-t)$, i.e., F is symmetric with centering constant $b/2$. ■

ACKNOWLEDGMENTS

I thank C. Burbeck, J. Kaiwi, and R. Saxon for assistance, and the University of California, Irvine, Focused Research Program on Higher Mental Processes for financial support.

REFERENCES

- BLOCK, H. D., & MARSCHAK, J. Random orderings and stochastic theories of responses. In I. Olkin, S. Ghurye, W. Hoeffding, W. Madow, and H. Mann (Eds.), *Contributions to probability and statistics*. Stanford: Stanford University Press, 1970.
- BURKE, C. J., & ZINNES, J. L. A paired comparison of pair comparisons. *Journal of Mathematical Psychology*, 1965, 2, 53-76.
- COHEN, M. A. Random utility systems—the infinite case. *Journal of Mathematical Psychology*, 1980, 22, 1-23.
- CORBIN, R., & MARLEY, A. A. J. Random utility models with equality: an apparent but not actual generalization of random utility models. *Journal of Mathematical Psychology*, 1974, 11, 174-293.
- FALMAGNE, J. C. A representation theorem for finite random scale systems. *Journal of Mathematical Psychology*, 1978, 18, 52-72.

- LUCE, R. D. *Individual choice behavior*. New York: Wiley, 1959.
- LUCE, R. D. The Choice Axiom after twenty years. *Journal of Mathematical Psychology*, 1977, **15**, 3, 215-233.
- LUCE, R. D., & SUPPES, P. Preference, utility, and subjective probability. In R. D. Luce, R. A. Bush, and E. Galanter (Eds.), *Handbook of Mathematical Psychology*, III. New York: Wiley, 1965.
- MARLEY, A. A. J. Some probabilistic models of simple choice and ranking. *Journal of Mathematical Psychology*, 1968, **5**, 333-359.
- MOSZNER, Z. On the equivalence of Thurstone models. *Journal of Mathematical Psychology*, 1978, **17**, 263-265.
- ROCKWELL, C., & YELLOTT, J. I. JR. A note on equivalent Thurstone models. *Journal of Mathematical Psychology*, 1979, **19**, 1, 61-64.
- STRAUSS, D. Some results on random utility models. *Journal of Mathematical Psychology*, 1979, **20**, 1, 35-52.
- THORSEN, S., & STEVER, J. Classes of models for selected axiomatic theories of choice. *Journal of Mathematical Psychology*, 1974, **11**, 15-32.
- THURSTONE, L. L. A law of comparative judgment. *Psychological Review*, 1927, **34**, 273-286.
- YELLOTT, J. I., JR. The relationship between Thurstone's, Luce's and Dawkin's models for pair comparisons. Paper given at annual Mathematical Psychology Meeting, Miami, September 1970.
- YELLOTT, J. I., JR. The relationship between Luce's Choice Axiom, Thurstone's theory of comparative judgment, and the double exponential distribution. *Journal of Mathematical Psychology*, 1977, **15**, 109-144.
- YELLOTT, J. I., JR., On a functional equation related to Thurstone models. *Journal of Mathematical Psychology*, 1978, **17**, 266-270.

RECEIVED: June 20, 1979