PSYCHOMETRIKA—VOL. 27, NO. 2
JUNE, 1962

AN OBSERVABLE PROPERTY EQUIVALENT TO A CHOICE MODEL FOR DISCRIMINATION EXPERIMENTS*

R. DUNCAN LUCE

UNIVERSITY OF PENNSYLVANIA

The choice model considered by Luce in an earlier work is stated in terms of unobservable parameters. In this paper a consequence of the model, involving only observables, is shown to be equivalent to the model.

In this note we prove that a trivial consequence of the choice model described in [1] for discrimination experiments is in fact equivalent to the model. This is of interest mainly because the property involves only observables, whereas the model itself is stated in terms of unobservable stimulus and response bias parameters.

The Experimental Situation

Abstractly, a general discrimination experiment can be described in terms of three sets and a function. The first set, $\mathcal{S}$, consists of the "elementary" *stimuli* selected and presented in various combinations to the subject by the experimenter. For example, in a loudness discrimination study $\mathcal{S}$ might consist of the 60, 62, 64, and 66 db 1000 cps tones. Of course, $\mathcal{S}$ is finite.

The second set, S, is a subset of $\mathcal{S} \times \mathcal{S} \times \cdots \times \mathcal{S}$ (ν times), and its elements we call *stimulus presentations*. Thus, a stimulus presentation is an ordered ν-tuple of the form $s = < s^1, s^2, \cdots, s^\eta, \cdots, s^\nu >$, where $s^\eta \in \mathcal{S}$. In the loudness example, time order is always used and, for $\nu = 3$, a typical presentation is the 64 db tone in the first time interval, the 60 db one in the second, and the 66 db one in the third. The subject is required to designate which of the three seems the "loudest." Although experimentalists are not entirely consistent in their usage, they often refer to our stimulus presentations simply as stimuli. In this paper we shall suppose $S = \mathcal{S} \times \mathcal{S} \times \cdots \times \mathcal{S}$, even though weaker assumptions would do.

The third set, R, consists of ν elements that we shall call responses. Continuing the example, they might be the three buttons, labeled "first," "second," and "third," which the subject presses to designate which of three successive tones is loudest. In some designs, a spatial rather than temporal

*This work was supported in part by National Science Foundation Grants NSF G-8864 and NSF G-17637 to the University of Pennsylvania.

identification of the stimuli is used. It is convenient to label the responses so that $R = \{1, 2, \cdots, i, \cdots, \nu\}$.

However the stimulus ordering and the labeling of the responses may be realized, the experimenter must establish a one-to-one relation, θ, between R and the several coordinates of the Cartesian product. This function assigns a meaning to the responses: by selecting response i when s is presented, the subject indicates that he believes coordinate $\theta(i)$ of s to contain the stimulus which satisfies the discriminative criterion. For example, his response designates which of the presented tones on a given trial he believes to have been the "loudest." It will be convenient to let s_i denote that element of $\mathcal{S}$ which happens to be located in the $\theta(i)$ coordinate of s, i.e., $s_i = s^{\theta(i)}$.

The Choice Model

For such an experimental structure, we assume that the (asymptotic) data arise from a trial independent random process generated by fixed conditional probabilities that response i is made when s is presented. Denote this by $p(i \mid s)$. Of course, we assume that

$$\sum_{i \varepsilon R} p(i \mid s) = 1.$$

The *choice model* stated in [1]—it is only described in detail for $\nu = 3$ but it is readily generalized—postulates two numerical ratio scales v and b.

$$v : \mathcal{S} \rightarrow \text{positive real numbers},$$

$$b : R \rightarrow \text{positive real numbers},$$

such that for any $p(i \mid s) \neq 0$ or 1,

$$p(i \mid s) = \frac{v(s_i)b(i)}{\sum_{j \varepsilon R} v(s_j)b(j)}.$$

The v's are interpreted as stimulus parameters—when they are viewed as a function of a physical measure of the stimuli they form part of a psychophysical scale—and the b's, as response biases. The argument in [1] leading to this model will not be repeated here.

By direct substitution, it is easy to see that this model implies the following.

PROPERTY *A*. *For any $r, s \varepsilon S$ and any $i, j, k \varepsilon R$ such that $r_i = s_j$,*

$$\frac{p(i \mid r)p(k \mid s)}{p(k \mid r)p(j \mid s)} = F(r_k, s_k, i, j).$$

Specifically,

$$F(r_k, s_k, i, j) = \frac{v(s_k)b(i)}{v(r_k)b(j)}.$$

Thus, if the choice model is correct, this particular combination of probabilities is not explicitly dependent upon the response k nor upon any of the stimuli comprising r and s save those which correspond to response k under the function θ. Put another way, if r' and s' are two other presentations such that $r'_i = s'_j$, $r'_k = r_k$, and $s'_k = s_k$, then the following testable relation among observables holds:

$$\frac{p(i \mid r)p(k \mid s)}{p(k \mid r)p(j \mid s)} = \frac{p(i \mid r')p(k \mid s')}{p(k \mid r')p(j \mid s')}.$$

The Result

THEOREM 1. *The (discrimination) choice model is equivalent to property A.*

PROOF. Because we already know that the model implies property A, we need only establish the converse. First, we show that $F(\rho, \rho, i, j) = g(i, j)$, independent of $\rho \in \mathcal{S}$. Choose a presentation r for which $r_i = r_j = \rho$ and $r_k = \sigma$, where $\rho, \sigma \in \mathcal{S}$; then by property A

$$F(\rho, \rho, i, j) = \frac{p(i \mid r)p(i \mid r)}{p(i \mid r)p(j \mid r)} = \frac{p(i \mid r)p(k \mid r)}{p(k \mid r)p(j \mid r)} = F(\sigma, \sigma, i, j).$$

To show $F(\rho, \sigma, i, i) = f(\rho, \sigma)$, independent of i, choose $r, s, t, u \in S$ for which $r_i = s_i = t_j = u_j$, $r_k = t_k = \rho$, and $s_k = u_k = \sigma$. Then

$$\begin{aligned}\frac{F(\rho, \sigma, i, i)}{F(\rho, \sigma, j, j)} &= \frac{p(i \mid r)p(k \mid s)}{p(k \mid r)p(i \mid s)} \Big/ \frac{p(j \mid t)p(k \mid u)}{p(k \mid t)p(j \mid u)} \\ &= \frac{p(i \mid r)p(k \mid t)}{p(k \mid r)p(j \mid t)} \Big/ \frac{p(i \mid s)p(k \mid u)}{p(k \mid s)p(j \mid u)} \\ &= \frac{F(\rho, \rho, i, j)}{F(\sigma, \sigma, i, j)} \\ &= 1.\end{aligned}$$

It follows immediately that $F(\rho, \rho, i, i) = 1$, so $f(\rho, \rho) = g(i, i) = 1$.

Now we establish that $F(\rho, \sigma, i, j) = f(\rho, \sigma)g(i, j)$. Choose any r, s, t, and u for which $r_i = s_i = t_i = u_j = \tau$, $r_k = \rho$, and $s_k = t_k = u_k = \sigma$, then

$$\begin{aligned}f(\rho, \sigma)g(i, j) &= F(\rho, \sigma, i, i)F(\sigma, \sigma, i, j) \\ &= \frac{p(i \mid r)p(k \mid s)p(i \mid t)p(k \mid u)}{p(k \mid r)p(i \mid s)p(k \mid t)p(j \mid u)} \\ &= \frac{p(i \mid r)}{p(k \mid r)} F(\tau, \tau, k, k) \frac{p(k \mid u)}{p(j \mid u)} \\ &= F(\rho, \sigma, i, j).\end{aligned}$$

Next we prove that $f(\rho, \sigma)f(\sigma, \tau) = f(\rho, \tau)$. Choose any r, s, t, and u

for which $r_i = s_i = t_j = u_j$, $r_k = \rho$, $s_k = t_k = \sigma$, and $u_k = \tau$, then

$$\begin{aligned} f(\rho, \sigma)f(\sigma, \tau) &= F(\rho, \sigma, i, i)F(\sigma, \tau, j, j) \\ &= \frac{p(i \mid r)p(k \mid s)p(j \mid t)p(k \mid u)}{p(k \mid r)p(i \mid s)p(k \mid t)p(j \mid u)} \\ &= \frac{p(i \mid r)}{p(k \mid r)} F(\sigma, \sigma, j, i) \frac{p(k \mid u)}{p(j \mid u)} \\ &= F(\rho, \tau, i, j)g(j, i) \\ &= f(\rho, \tau)g(i, j)g(j, i). \end{aligned}$$

But by choosing a new r and s for which $r_i = s_j = \rho$ and $r_k = s_k = \sigma$, we see that

$$\begin{aligned} g(i, j)g(j, i) &= F(\sigma, \sigma, i, j)F(\sigma, \sigma, j, i) \\ &= \frac{p(i \mid r)p(k \mid s)p(j \mid s)p(k \mid r)}{p(k \mid r)p(j \mid s)p(k \mid s)p(i \mid r)} \\ &= 1. \end{aligned}$$

To show the parallel equation for g, choose any r, s, t, and u for which all the relevant stimuli are ρ, then

$$\begin{aligned} g(i, j)g(j, k) &= F(\rho, \rho, i, j)F(\rho, \rho, j, k) \\ &= \frac{p(i \mid r)p(l \mid s)p(j \mid t)p(l \mid u)}{p(l \mid r)p(j \mid s)p(l \mid t)p(k \mid u)} \\ &= F(\rho, \rho, i, k)F(\rho, \rho, l, l) \\ &= g(i, k). \end{aligned}$$

For some fixed $\rho^* \varepsilon \mathcal{S}$, define $v(\rho) = f(\rho^*, \rho)$. By what we have just proved, $v(\rho^*) = f(\rho^*, \rho^*) = 1$ and $f(\rho, \sigma) = f(\rho^*, \sigma)/f(\rho^*, \rho) = v(\sigma)/v(\rho)$. Similarly for some $i^* \varepsilon R$, define $b(i) = g(i, i^*)$. Then $b(i^*) = 1$ and $g(i, j) = b(i)/b(j)$.

Finally, we show that these two scales satisfy the choice model. Let s, i, and j be given and choose any r such that $r_j = s_i$, then

$$\begin{aligned} \frac{p(i \mid s)}{p(j \mid s)} &= \frac{p(i \mid s)p(j \mid r)}{p(j \mid s)p(j \mid r)} \\ &= f(s_j, r_j)g(i, j) \\ &= \frac{v(r_j)b(i)}{v(s_j)b(j)} \\ &= \frac{v(s_i)b(i)}{v(s_j)b(j)}. \end{aligned}$$

Summing over all $j \varepsilon R$ yields the choice model.

Assuming that property A holds, the proof of Theorem 1 gives an explicit way to estimate the two scales of the choice model. For any r and s such that $r_i = s_i$, $r_k = \rho^*$, and $s_k = \rho$,

$$\begin{aligned} v(\rho) &= f(\rho^*, \rho) \\ &= \frac{p(i \mid r)\, p(k \mid s)}{p(k \mid r)\, p(i \mid s)}. \end{aligned}$$

With fallible data, each such r and s pair generally yields a different estimate of $v(\rho)$. Rather than take an arithmetic average, it seems more appropriate in this model to take the appropriate root of the product of all such estimates (different from 0 and ∞) as the "average" estimate. Similarly, for any r and s such that $r_i = s_{i*}$ and $r_k = s_k = \rho$,

$$\begin{aligned} b(i) &= g(i, i^*) \\ &= \frac{p(i \mid r)}{p(k \mid r)} \frac{p(k \mid s)}{p(i^* \mid s)}. \end{aligned}$$

The same sort of average estimate seems appropriate, although these estimates have no known statistical properties.

The main significance of Theorem 1 is that the "observable" statements embodied in property A are sufficient to imply the choice model; any other observable relations are logical consequences of these.

The Unbiased Case

Consider now any presentations r and s which are permutations of a fixed subset of $\mathcal{S}$. We say that the *order of presentation is irrelevant* provided that $r_i = s_j$ implies $p(i \mid r) = p(j \mid s)$.

THEOREM 2. *If the choice model holds, the order of presentation is irrelevant if and only if* $b(i) = 1$ *for all* $i \in R$.

PROOF. It is obvious that if $b(i) = 1$ for all i, then the order of presentation is irrelevant. Conversely, for any i and $j \in R$, choose permutations r and s such that $r_i = s_j$ and $r_k = s_k$, then if the order is irrelevant

$$\begin{aligned} g(i, j) &= F(r_k, r_k, i, j) \\ &= \frac{p(i \mid r) p(k \mid s)}{p(k \mid r) p(j \mid s)} \\ &= 1. \end{aligned}$$

Thus, $b(i) = g(i, i^*) = 1$.

REFERENCES

[1] Luce, R. D. *Individual choice behavior: a theoretical analysis.* New York: Wiley, 1959.

Manuscript received 10/8/60
Revised manuscript received 7/6/61