Measuring laterality effects in dichotic listening

Bruno H. Repp

Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511
(Received 3 March 1977; revised 5 May 1977)

This paper discusses methodological issues and problems related to measuring laterality effects in dichotic listening. Section I describes the standard dichotic two-response paradigm, as well as a number of indices of the degree of ear advantage proposed in the literature. The numerical range of most of these indices is constrained by performance level; only one particular index avoids these constraints. However, this does not make this index necessarily the optimal one. First, a correction for guessing is proposed. Then, analogies to signal-detection theory are discussed, as well as theoretical and empirical criteria for choosing the correct index of laterality. While this index is still unknown, the index called here e_g is proposed as the best solution at the present state of knowledge. Section II discusses the phenomenon of dichotic fusion and the dichotic single-response paradigm, which offers many methodological advantages over the two-response paradigm. Section III discusses the factors of ear dominance and stimulus dominance in the perception of fused stimuli. An index of ear dominance is derived by again taking advantage of analogies to signal-detection theory. In Sec. IV, a number of remaining problems are discussed: stimulus intelligibility, guessing and selective attention, blend responses, test reliability, validity, and homogeneity.

PACS numbers: 43.70.Dn, 43.66.Rq, 43.66.Yw

INTRODUCTION

Since Kimura's (1961) demonstration of an average right-ear advantage (REA) in the recognition of dichotic verbal stimuli, many researchers have used dichotic listening tasks to measure hemispheric dominance for language. Kimura's hypothesis that hemispheric dominance for language is the major factor underlying the ear asymmetries has been almost universally accepted. While some studies have been content with diagnosing the mere direction of the average ear advantage (left or right) and testing its significance, many recent studies have attempted to compare different individuals, different tests, or different experimental conditions with respect to the observed magnitude of the ear asymmetry. Underlying these attempts has been the belief that cerebal lateralization, like handedness, is a matter of degree and can be measured on a continuous scale (Zangwill, 1960; Shankweiler and Studdert-Kennedy, 1975).

In order to yield meaningful and reliable measurements, dichotic testing must meet certain formal and methodological requirements which have been given relatively little attention in the past. If dichotic listening tasks are used as instruments to measure the degree of hemispheric dominance for language, they must satisfy the same high standards of construction, procedure, and scoring as any other psychological test. These standards may be derived from methodologically oriented research in the laboratory, from theoretical analyses of the task situation, and from general test-theoretical principles. Many of these requirements are not sufficiently met by dichotic tests as they are now widely used.

The present paper summarizes the issues that must be dealt with in constructing an effective dichotic test for the purpose of measuring hemispheric dominance. The dichotic listening situation is remarkably complex. In the discussion that follows, I provide some suggestions but point out many problems that need further investigation or have not been dealt with at all in the past. Although the discussion is restricted to dichotic listening, many of the issues should also apply to other situations in which lateral asymmetries are measured (e.g., visual hemifield differences, binocular rivalry, or ocular dominance), and therefore may be of interest to a wider audience.

One focus of the present discussion is the problem of choosing a numerical index of the ear advantage. This problem is fundamental to the measurement of lateralization; unless it is solved, no meaningful comparisons between subjects, tests, or experimental conditions are possible. Section I—which heavily relies on earlier discussions by Halwes (1969) and Marshall, Caplan, and Holmes (1975)—discusses a number of indices that have been proposed and used in the past in conjunction with the dichotic two-response paradigm (which requires the listener to identify both stimuli in a dichotic pair). All but one of these indices fail to take into account the constraints imposed by performance level on the range of differences between the scores for the two ears. In addition, none of them corrects for guessing, despite the fact that most dichotic studies use only a few different stimuli, which results in substantial guessing probabilities. An index is described that takes both performance level and guessing into account. However, the correct index must be based on a correct theory and empirical evidence of how scores for the two ears change with performance level and how guessing operates, and this theoretical and empirical basis is not available at present. The index proposed here is merely based on plausible assumptions; the question whether it is *the* correct index remains open.

The second focus of the present paper is finding ways of simplifying dichotic testing and circumventing some of the problems encountered in the standard two-response paradigm. Sections II and III describe an approach to dichotic listening that in many ways seems simpler than the two-response paradigm. This method, which requires only a single response to each dichotic stimulus, relies on the phenomenon of dichotic (or binaural) fusion. Section II briefly discusses the factors that make two dichotic stimuli fuse more or less completely into a single perceived stimulus, as well as the methodological consequences of such fusion. Section III derives an index of the ear advantage for the single-response paradigm. Here, the phenomenon of stimulus dominance (perceptual dominance of one stimulus over the other in a fused dichotic pair) exerts constraints on the ear score difference similar to that exerted by performance level in the two-response paradigm. However, these constraints can be dealt with and actually become a crucial factor in deriving an unbiased index of the ear advantage.

Section IV is devoted to a survey of additional topics and problems in dichotic testing: stimulus intelligibility, selective attention, blend responses, test reliability, homogeneity, and validity. Since the concern of this paper is exclusively methodological, any discussion of the physiological factors that may underly dichotic ear advantages is avoided. The aim is to develop methods for measuring the dichotic ear advantage with maximum precision. Before we can attempt to answer the more fundamental questions about the structures and processes underlying the ear asymmetry, we must be able to obtain valid and reliable measurements from dichotic tests. There is much room for improvement in existing methods with respect to this goal.

I. LATERALITY INDICES FOR THE TWO-RESPONSE PARADIGM

A. Method

In the two-response paradigm, two different stimuli are presented simultaneously to the two ears, and the subject is asked to identify both, typically without any constraint on the order of report. The two responses must be different from each other, and guessing is encouraged. This is the standard situation that will be considered in this section.

The results of a standard two-response test may be summarized in a 2×2 table, as shown in Table I. The responses are scored as correct (i.e., identical with one of the stimuli) or incorrect without regard to order. The proportions of correct and incorrect responses are calculated separately for each ear, so that the row sums in Table I are equal to 1.0.

The *overall performance level* is defined as the average proportion of correct responses per ear,

TABLE I. The data structure in the two-response paradigm.

		Responses		
		Correct	Incorrect	
Channels	LE	P_L	$1-P_L$	1.00
	RE	P_R	$1-P_R$	1.00
		P_L+P_R	$2-P_L-P_R$	2.00

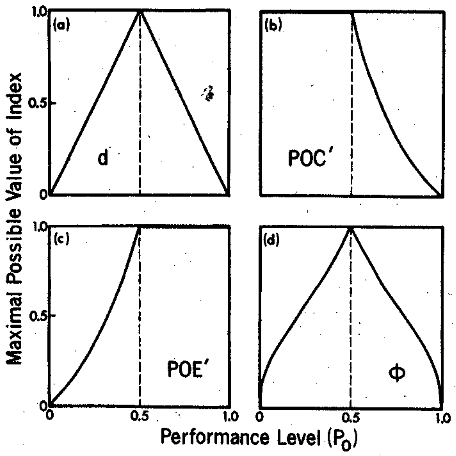

FIG. 1. The numerical ranges of four indices of the ear advantage as a function of performance level.

$$P_0=\tfrac{1}{2}(P_R+P_L) \quad (1)$$

B. The simple difference score (*d*)

The simplest index of the ear advantage is the difference between the proportions of correct responses for the two ears,

$$d=P_R-P_L \quad (2)$$

The vast majority of dichotic listening studies have reported the ear advantage as *d*. (There is no commonly accepted name of the index; it is called *d* here simply for notational convenience.) Although useful as a descriptive statistic, *d* has severe limitations when the results of different subjects, different tests, or different experimental conditions are to be compared. These limitations arise from the constraint imposed on differences between proportions by their absolute size—a fact that is often neglected and therefore constitutes one of the primary fallacies of descriptive statistics. In the context of measuring laterality effects, Halwes (1969) was the first to point out that the overall performance level P_0 sets an upper limit to *d*,

$$d_{max}=2P_0 \quad \text{if } 0.0\leq P_0\leq 0.5 ,$$
$$d_{max}=2(1-P_0) \quad \text{if } 0.5\leq P_0\leq 1.0 , \quad (3)$$

where d_{max} is the maximal value that *d* can assume at a given level of P_0, and $d_{max}=-d_{min}$, the corresponding minimal (i.e., maximal negative) value. Figure 1(a) shows the triangular function respresented by Eq. (3).

Thus, *d* indices of different subjects, tests, or experimental conditions are not directly comparable unless the respective performance levels are equal. Since, in general, performance levels are not constant from one subject (test, condition) to another, comparisons of *d* indices are almost certainly invalid. Many studies in the past have neglected this quite elementary limitation of simple difference scores and, consequently, some of these studies may have reached faulty conclusions.[1] Nevertheless, it must be conceded that there is a theoretical possibility that direct comparisons of *d* are valid, after all—i.e., that *d* is the correct index to use—, although this seems very unlikely (see Sec. I E).

Hopefully, empirical evidence will become available in the future to decide this issue objectively.

C. "Correcting" for the constraints of performance level

Several authors became aware of the limitations of d and proposed alternative indices of the ear advantage that were subsequently used by others. All of these indices were intended to provide a measure of the ear advantage that is independent of performance level, both theoretically and empirically. Only one of the indices discussed here seems to achieve this aim, but it will be argued that this index is still not a final solution to the problem of finding the optimal index of the ear advantage.

POC (percentage of correct [responses]) and POE (percentage of errors) are two alternative indices suggested by Harshman and Krashen (1972). They are defined as

$$\mathrm{POC} = P_R/(P_R + P_L) \quad (4)$$

and

$$\mathrm{POE} = (1 - P_L)/(2 - P_R - P_L) \ . \quad (5)$$

These indices range from 0 (perfect LEA) to 1 (perfect REA); an index of 0.5 means no ear advantage. For those who, like the present author, prefer a scale ranging from -1 (perfect LEA) to $+1$ (perfect REA)—and this is entirely a matter of personal choice—corresponding POC′ and POE′ indices are obtained by a simple linear transformation of POC and POE[2]:

$$\mathrm{POC}' = 2\mathrm{POC} - 1 = (P_R - P_L)/(P_R + P_L) \ , \quad (6)$$

$$\mathrm{POE}' = 2\mathrm{POE} - 1 = (P_R - P_L)/(2 - P_R - P_L) \ . \quad (7)$$

The limitations of POC and POE as a function of performance level have recently been competently discussed by Marshall *et al.* (1975). The analogous limitations of POC′ and POE′ are illustrated in Figs. 1(b) and (c). In formal terms, we obtain from Eqs. (3), (6), and (7),

$$\mathrm{POC}'_{\max} = 1 \quad \text{if } 0.0 \le P_0 \le 0.5 \ ,$$
$$\mathrm{POC}'_{\max} = (1 - P_0)/P_0 \quad \text{if } 0.5 \le P_0 \le 1.0 \ , \quad (8)$$

$$\mathrm{POE}'_{\max} = P_0/(1 - P_0) \quad \text{if } 0.0 \le P_0 \le 0.5 \ ,$$
$$\mathrm{POE}'_{\max} = 1 \quad \text{if } 0.5 \le P_0 \le 1.0 \ . \quad (9)$$

Thus, it is evident that the range of POC′ is unconstrained at low performance levels and the range of POE′ is unconstrained at high performance levels, but where one index is unconstrained the other is severely limited by performance level. The same is true for POC and POE. Harshman and Krashen (1972) preferred POE over POC after empirically demonstrating a high positive correlation between P_0 and POC but a low correlation between P_0 and POE, as computed over a number of studies in the literature. This finding can be explained by the fact that high performance levels are more commonly encountered in dichotic studies than low performance levels, so that the majority of the reported scores fell in the region where $\mathrm{POE}_{\max}$ rather than $\mathrm{POC}_{\max}$ is independent of performance level.

A quite different approach was taken by Kuhn (1973) who proposed an existing statistical index, the ϕ coefficient, as the solution to the performance level problem, which was readily accepted by several other authors. However, Levy (in press) has presented mathematical proof and empirical evidence that the ϕ coefficient does depend on performance level. The theoretical argument can be made in simplified form by pointing out the relationship between ϕ and POC′ and POE′:

$$\phi = (P_R - P_L)/[(P_R + P_L)(2 - P_R - P_L)]^{1/2}$$
$$= [(\mathrm{POC}')(\mathrm{POE}')]^{1/2} \ . \quad (10)$$

Then, from Eqs. (8), (9), and (10),

$$\phi_{\max} = [P_0/(1 - P_0)]^{1/2} \quad \text{if } 0.0 \le P_0 \le 0.5 \ ,$$
$$\phi_{\max} = [(1 - P_0)/P_0]^{1/2} \quad \text{if } 0.5 \le P_0 \le 1.0 \ . \quad (11)$$

Thus, $\phi_{\max}$, much like $d_{\max}$, is constrained by P_0 at all performance levels except 0.5. This is illustrated in Fig. 1(d).

Being a conjunction of POC′ and POE′—viz., their geometric mean—ϕ combines the constraints of these two indices. The most obvious solution is a disjunctive use of POC′ and POE′ that takes advantage of the fact that each is unconstrained in one half of the range of P_0. Thus,

$$e = \mathrm{POC}' = (P_R - P_L)/(P_R + P_L) \quad \text{if } 0.0 \le P_0 \le 0.5 \ ,$$
$$e = \mathrm{POE}' = (P_R - P_L)/(2 - P_R - P_L) \quad \text{if } 0.5 \le P_0 \le 1.0 \ . \quad (12)$$

Since $e = d/d_{\max}$ [cf. Eqs. (2) and (3)], $e_{\max} = 1$ and thus is completely independent of P_0.[3] The idea to express the observed ear difference as a proportion of the maximally possible ear difference at a given performance level was first conceived by Halwes (1969) and, more recently and apparently independently, by Marshall *et al.* (1975) who called their index f. The solution seems straightforward—it is a simple multiplicative rescaling of d to fit its restricted range.

It should be noted that the range of e is independent of P_0 but not its variability: e is least variable (i.e., most reliable) at intermediate levels of P_0, and it gets unreliable when P_0 approaches extremely low or high levels. In addition, the variance of e is likely to vary with the magnitude of e itself. It is important not to confuse the magnitude (estimated mean) of an index with its variance; these are separate statistical parameters, and only the former is expressed by an ear advantage index. Thus, the magnitude of an index is not a direct measure of its significance, which requires a separate test.

Despite its apparent independence of P_0, e is not necessarily the optimal index. The kind of theoretical and empirical support that is needed to determine the correct index will be discussed in Sec. I E (see also Marshall *et al.*, 1975). At this point, let us consider a more obvious shortcoming of the e index (and of all other indices proposed, for that matter): its failure to correct for guessing. Strangely enough, a correction for guessing has never been considered in the past,

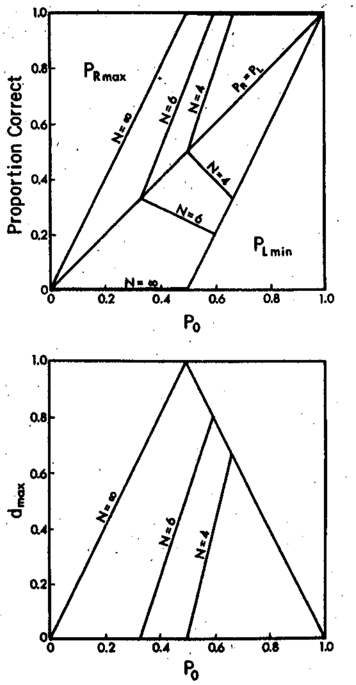

FIG. 2. Maximal and minimal ear scores (upper panel) and their difference (lower panel) as a function of performance level, for three levels of guessing probability.

although it is obvious that guessing plays a substantial role in most dichotic experiments. In the next section, an appropriate correction for guessing is derived.

D. Correction for guessing

In order to deal with the guessing problem, we need to consider the scores for each ear, not just their difference d, as a function of P_0. This is illustrated in Fig. 2(a). The diagonal line labeled $P_R = P_L$ is the case of no ear advantage ($d=0$). In this case, $P_R = P_L = P_0$, regardless of the guessing probability. At the other extreme, consider the maximal and minimal possible ear scores, $P_{R\,max}$ and $P_{L\,min}$, as a function of P_0. (We assume here, without loss of generality, that the right ear is the dominant ear; the corresponding results for left-ear advantage are obtained by interchanging the R and L subscripts.) Let us first assume that N, the number of stimuli, equals infinity, so that the guessing probability is zero. Then the lowest possible performance level, $P_{0\,min}$, is zero, and, of course, $P_{R\,max} = P_{L\,min} = 0$ if $P_{0\,min} = 0$. As P_0 increases, $P_{R\,max}$ increases linearly towards 1.0 while $P_{L\,min}$ remains at 0; consequently, $P_{R\,max} = 2P_0$ and $P_{L\,min} = 0$ for $0.0 \leq P_0 \leq 0.5$. At $P_0 = 0.5$, $P_{R\,max}$ reaches 1.0 and remains at this level while $P_{L\,min}$ begins to increase with P_0; consequently, $P_{R\,max} = 1.0$ and $P_{L\,min} = 2P_0 - 1$ for $0.5 \leq P_0 \leq 1.0$. Thus, the maximally divergent scores for the two ears are represented by the parallelogram labeled $N=\infty$ in Fig. 2(a). Of course, $P_{R\,max} - P_{L\,min} = d_{max}$, whose relation to P_0 is shown in Fig. 1(a) and again in Fig. 2(b) as the function labeled $N=\infty$.

Now consider the more realistic case of a nonzero guessing probability. Two typical cases, $N=6$ and $N=4$, are illustrated in Fig. 2(a). The lowest expected performance level for a given number of stimuli, $P_{0\,min}$, is found to be

$$P_{0\,min} = (N-1)\Big/\binom{N}{2} = 2/N\,. \tag{13}$$

This is the performance level that would be expected if the subject produced only completely random guesses, because $(N-1)$ of the possible $\binom{N}{2} = \frac{1}{2}N(N-1)$ combinations of two responses lead by chance to a correct response for one ear. Thus, $P_{R\,max} = P_{L\,min} = P_{0\,min}$ if $P_0 = P_{0\,min}$. From this minimum, $P_{R\,max}$ increases linearly towards 1.0 as P_0 increases while $P_{L\,min}$ remains at chance level. However, this chance level does not remain constant but depends on $P_{R\,max}$. At the point of maximal ear difference, where $P_{R\,max}$ reaches 1.0,

$$P_{L\,min} = 1/(N-1)\,, \tag{14}$$

which is the simple guessing probability for N stimuli. (It is not $1/N$ because the right-ear response must be different from the left-ear response.) In other words, at this point a hypothetical listener with the maximal possible ear difference always can identify the stimulus in the dominant ear but produces a random guess for the stimulus in the other ear. The maximal ear difference d_{max} at this point is

$$d_{max} = P_{R\,max} - P_{L\,min} = 1 - 1/(N-1) = (N-2)/(N-1), \tag{15}$$

which is the maximal expected ear difference for a given N. It occurs at a performance level of

$$P_0 = \tfrac{1}{2} + \tfrac{1}{2}(N-1) = N/2(N-1)\,. \tag{16}$$

From this point on, $P_{R\,max}$ remains at 1.0 and $P_{L\,min}$ increases with P_0. The complete functions relating $P_{R\,max}$ and $P_{L\,min}$ to P_0 are

$$\begin{aligned} P_{R\,max} &= [2/(N-2)][(N-1)P_0 - 1] \\ &\qquad \text{if } 2/N \leq P_0 \leq N/2(N-1)\,, \end{aligned} \tag{17}$$

$$\begin{aligned} P_{R\,max} &= 1 && \text{if } N/2(N-1) \leq P_0 \leq 1.0\,, \\ P_{L\,min} &= [2/(N-2)](1-P_0) && \text{if } 2/N \leq P_0 \leq N/2(N-1)\,, \\ P_{L\,min} &= 2P_0 - 1 && \text{if } N/2(N-1) \leq P_0 \leq 1.0\,. \end{aligned} \tag{18}$$

Figure 2(b) shows the corresponding relationship between d_{max} and P_0 for $N=\infty$, $N=6$, and $N=4$. For a finite N, this function is

$$\begin{aligned} d_{max} &= P_{R\,max} - P_{L\,min} \\ &= [2/(N-2)](NP_0 - 2) && \text{if } 2/N \leq P_0 \leq N/2(N-1)\,, \\ &= 2(1-P_0) && \text{if } N/2(N-1) \leq P_0 \leq 1.0\,. \end{aligned} \tag{19}$$

(The function for $N=\infty$ is given in Eq. (3).)

Now we define e_g—as we will call e with the correction for guessing—as

$$\begin{aligned} e_g &= d/d_{max} \\ &= (P_R - P_L)/[2(NP_0 - 2)/(N-2)] && \text{if } 2/N \leq P_0 \leq N/2(N-1)\,, \\ &= (P_R - P_L)/[2(1-P_0)] && \text{if } N/2(N-1) \leq P_0 \leq 1.0\,. \end{aligned} \tag{20}$$

Equation (20) shows that e_g is identical to e—and thus to POE′—in the upper range of performance levels. In other words, POE′ is unaffected by guessing probability and needs no correction. It is only in the lower range of performance, where POC′ applies, that a correction for guessing becomes necessary. Without it, the magnitudes of ear advantage at low performance levels would be seriously underestimated.

The correction for guessing just proposed is only a global and approximate solution. Ideally, such a correction should be based on a detailed model of perceptual and response processes in dichotic listening. At present, such a model does not exist. Repp (1977b) has considered a very simple probabilistic model that assumes that the listener either perceives a stimulus correctly or makes a random guess, independently for each ear. The *e* index based on the resulting estimates of the "true" probabilities of perceiving left- and right-ear stimuli is almost identical to e_g. However, the model is too simple to provide a complete account of the perception of dichotic stimuli; likewise, the correction for guessing may have to be refined in the future.

E. Isolaterality contours

The correct index of the ear advantage must fit both theoretical conceptions and empirical evidence. Halwes (1969) believed to have solved the theoretical problem by proposing an index (*e*) whose range is free of the constraints of performance level. Although *e* may very well be the correct index to use, Halwes' argument—expounded in Sec. I C—is not completely valid, as the following discussion will show. Marshall *et al.* (1975), who also proposed *e* as the perhaps best index, correctly stressed that different indices represent "psychological theories of how an *S* (subject) changes Rc and Lc [P_R and P_L] in achieving different overall accuracies" (p. 320).[4] In other words, performance level must be understood as a *consequence* of changes in right-ear and left-ear scores, and the concomitant constraints on the ranges of certain indices must be accepted if they are predicted by theories about the form of covariation of P_R and P_L. There is no such theory that explicitly postulates that the range of an ear advantage index must not be constrained in any region of performance.

However, among the infinite number of possible theories, there is one class of theories that lead precisely to this outcome, e.g., the theory underlying the *e* index. In order to clarify this point, consider the *isolaterality contours* assumed by different indices, i.e., by different theories of the ear advantage. Isolaterality contours connect points of equal underlying ear asymmetry at different levels of performance. In Fig. 1, these contours would be parallel horizontal lines within the limits of each index. It is more illuminating to represent these isolaterality contours in terms of P_R and P_L, as Marshall *et al.* (1975) have done. Figure 3 plots P_R against P_L, so that the isolaterality contours connect all pairs of scores (P_R, P_L) that are assumed to reflect the same underlying ear asymmetry. To simplify the exposition, the guessing probability is assumed to be zero; a nonzero guessing probability would have the effect of restricting the possible score combinations to a region in the upper right-hand corner of the unit square (or accuracy space, as Marshall *et al.* call it).

Figure 3 shows the isolaterality contours assumed by four theories: those associated with the indices *d*, POC′, POE′, and *e*. Note that the region above the positive diagonal represents REA's while the symmetric region below the positive diagonal represents LEA's. The isolaterality contours are shown only for REA's; those for LEA's are obtained by symmetric reflection around the positive diagonal. The *isoperformance contours*, which connect all pairs of scores (P_R, P_L) at the same $P_0 = \frac{1}{2}(P_R + P_L)$, are straight lines parallel to the negative diagonal in each case.

Figure 3 shows that only the *e* index provides a definite estimate of the magnitude of the ear advantage for every pair of scores. The other three indices depicted can give only a lower or upper bound on the ear advantage when one of the two ear scores is either at chance level or perfect, because these data points cannot be uniquely assigned to a particular isolaterality contour. Thus, for example, the fact that *d* cannot exceed 0.2 when $P_0 = 0.8$—due to the "constraint imposed by performance level on the range of the index," discussed in connection with Fig. 1—really means that, if *d* is the correct index (i.e., if the theory underlying *d* is correct), any true ear advantage of $d > 0.2$ cannot be measured at $P_0 = 0.8$. If the model underlying *d* happened to be correct, this disadvantage must be accepted; it cannot be taken as an *a priori* argument against the index theory. Similar arguments apply to POC′ and POE′.

From Fig. 3, the close analogy to signal detection theory is evident, which has also been pointed out by Marshall *et al.* (1975). P_L is formally analogous to the false alarm probability, and P_R to the hit probability in signal detection. Isolaterality contours correspond to receiver operating characteristic (ROC) functions, and isoperformance contours to isobias contours (cf. Green and Swets, 1966). The isolaterality contours assumed by the *e* index [Fig. 3(d)] are linear approximations to the ROC functions resulting from the standard signal detection model assuming underlying normal distributions with equal variance. This gives the *e* index some intuitive plausibility, in view of the success of the standard signal detection model in many different situations. However, whether it is also a correct model of

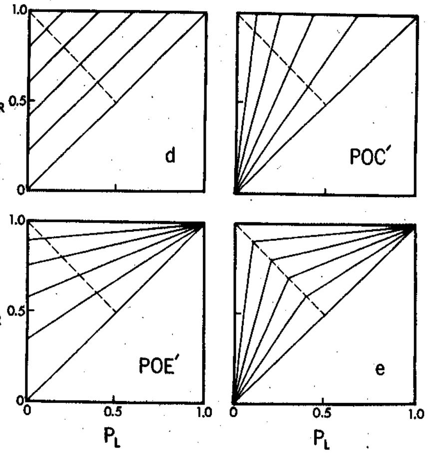

FIG. 3. Isolaterality contours assumed by the theories underlying four indices of the ear advantage.

the dichotic ear advantage remains to be proven. In the absence of stronger theoretical or empirical support, the alternative models underlying d, POC′, or POE′ cannot be ruled out. The POE′ model, for example, corresponds to a "high-threshold" model in terms of signal detection theory, which has been found to apply in certain situations (Green and Swets, 1966). There is an infinity of other possible models; those depicted in Fig. 3 are merely the extreme cases.

In addition to the intuitive appeal of e, its underlying assumptions may be plausibly conceptualized as follows. Assume that differences in performance level reflect different levels of noise in the perceptual-auditory system of listeners. Further, assume that, as the internal noise level is reduced from very high to very low, P_R and P_L increase independently of each other in the form of two psychometric (ogive) functions. The separation between these functions equals the true ear asymmetry and may be expressed in terms of the signal-detection statistic d'. This simple conception is identical with the standard signal detection model, so that e—whose isolaterality contours are a good approximation to the standard model—would be the correct index (if not d' itself is chosen as the index, which certainly is an option). Again, however, this argument has only intuitive plausibility at present. A decision between different models will require empirical evidence in favor of one or the other.

Unfortunately, empirical tests of the models are difficult. Marshall *et al.* (1975) have pointed out that, in analogy to signal detection, it would be necessary to vary performance level (which corresponds to the criterion or cutoff point) in a number of steps while holding the underlying ear asymmetry (which corresponds to sensitivity) constant. This would generate points on the same ROC function whose shape could then be determined. There are both theoretical and practical problems with this approach. The most obvious technique would be to employ masking noise or some other form of distortion to vary performance level within a single subject, but it is not clear whether this variation would be equivalent to the hypothetical variations in internal noise level that cause variations in P_0 *between* subjects, despite high monaural intelligibility of the stimuli. Cullen, Thompson, Hughes, Berlin, and Samson (1974) have varied signal-to-noise ratio in a dichotic two-response paradigm, but their results are irregular and permit no conclusion. A practical problem is that ear advantages tend to be rather small and highly variable, so that an enormous amount of data would be necessary to distinguish different shapes of ROC functions in the vicinity of the positive diagonal.

Halwes (1969) took the more global empirical approach of taking the average ear differences obtained for different groups of subjects in a number of different experiments and plotting them as a function of the variations in average performance level between the experiments. When the average ear advantages were expressed in terms of e, they turned out to be strikingly independent of performance level, which, at the time, provided impressive empirical support for the e index. Unfortunately, this result does not hold up in the light of more recent data. The present author has surveyed a large number of dichotic studies conducted since 1969 and found large variations in the magnitudes of ear advantages from study to study, regardless of performance level, so that no clear conclusion emerges from the data in the literature.

Yet another way of testing the assumptions underlying the e index would be to conduct an analysis of individual stimulus pairs. A large amount of data would have to be collected for this purpose. If individual stimulus pairs vary in their "performance level," but not in the underlying ear asymmetry, then the data points (P_R plotted against P_L) for all individual stimulus pairs should lie on the same ROC function. This approach is analogous to that discussed in Sec. III for the single-response paradigm, and it is certainly worth investigating. However, it is not clear whether individual stimulus pairs vary more than randomly in performance level (except for the "feature-sharing effect" discussed in Sec. IV); performance level has so far been considered a characteristic of the listener, not of the stimuli. More detailed investigations of the dichotic competition between individual stimuli are needed.

Thus, although e has the advantage of being the intuitively most satisfying index, other indices and their corresponding models cannot be ruled out completely at present. It is recommended, however, that e_g [i.e., e with the correction for guessing—Eq. (20)] be adopted as an index as long as there is no evidence that speaks against its use. In the remainder of this paper, a simpler approach to measuring the ear advantage will be described that, despite many analogies, avoids some of the problems inherent in the two-response paradigm. Some of these problems will become clear as the discussion proceeds (see especially Sec. IV). Considering the complexity of the two-response paradigm, it may be time to look for alternative methods that perhaps achieve the same goal with fewer complications.

II. DICHOTIC FUSION AND THE SINGLE-RESPONSE PARADIGM

A. Dichotic fusion

When two sounds are presented simultaneously to the two ears, they are not always perceived as two separate events. Often they fuse into a single sound image, i.e., binaural (dichotic) fusion occurs. The fusion mechanism tolerates a certain amount of spectral discrepancy. For example, dichotic sinusoids within a certain critical frequency range (the "binaural critical band") are heard as a single tone, although it may "beat" when low frequencies are involved (Odenthal, 1963; Perrott and Barry, 1969; Van den Brink, Sintnicolaas, and Van Stam, 1976). The fused tone is heard at a frequency intermediate between the two dichotic frequencies (Odenthal, 1963). Of special importance is the finding that two different tones that normally would not fuse can be made to fuse by imposing the same low-frequency modulation onto them (Leakey, Sayers, and Cherry, 1958, Tobias, 1972). In general, it seems that complex auditory signals with similar waveform envelopes fuse despite considerable differences in microstructure.

This result is important in the dichotic fusion of speech sounds. The waveform envelope of a speech signal is determined by its low-frequency components (primarily the fundamental frequency), while the higher formants constitute the microstructure. Two different isolated formants presented dichotically at the same fundamental frequency fuse into a single sound, while two formants with the same center frequency but with different fundamental frequencies are heard as separate sounds (Broadbent and Ladefoged, 1957). Thus, a speech signal may be "split" by filtering it into nonoverlapping low- and high-frequency bands which, if presented simultaneously to the two ears, are heard as a single source resembling the original (Broadbent, 1955; Franklin, 1969). Several recent studies have employed the related "split-formant technique" with synthetic speech, where some formants are presented to one ear and the remaining formants to the other ear (Rand, 1974; Nye, Nearey, and Rand, 1974; Nearey and Levitt, 1974; Haggard, 1975). Cutting (1976), in his recent classification of dichotic fusion phenomena, called this "spectral fusion."

Dichotic fusion is not limited to the case where parts of a speech signal fuse to reconstitute the original whole stimulus. Even if two different complete utterances are presented, the perceptual result may be a single fused stimulus, provided that the two dichotic stimuli have identical fundamental frequencies. The fused percept may resemble one or the other component, or it may be a hybrid (see Cutting, 1976). In assessing dichotic ear differences, it is important to know whether some or all of the stimuli fuse. Ideally, the experimenter should be able to control this property of the stimuli.

The verbal materials used in dichotic listening studies may be roughly classified into three groups:

(1) Words, digits, and other larger-sized verbal units. Typically, they are natural speech and acoustically heterogeneous, so that the waveforms in the two ears show little correspondence. Therefore, they tend not to fuse.[5]

(2) Natural-speech nonsense syllables, which have been used extensively in recent research (e.g., Studdert-Kennedy and Shankweiler, 1970; Berlin, Lowe-Bell, Cullen, Thompson, and Loovis, 1973; Cullen *et al.*, 1974). The typical set is /ba/, /da/, /ga/, /pa/, /ta/, and /ka/, spoken by the same voice. Some of the dichotic pairs formed from these syllables may fuse into a single syllable if they are spectrally similar and properly synchronized; this will depend on the particular stimuli and recording procedures used. Apart from temporal alignment, however, the experimenter has little control over fusion. Tests of this kind often contain fused and unfused pairs mixed together, which is a methodological disadvantage.

(3) Synthetic syllables (e.g., Halwes, 1969; Shankweiler and Studdert-Kennedy, 1967, 1975). As with any other stimuli, it depends on their spectral similarity (most of all on their fundamental frequencies) whether they do or do not fuse. However, the important advantage of synthetic syllables is that their acoustic properties—and, hence, their tendency to fuse—are under the control of the experimenter. Thus it is possible to construct homogeneous tests that contain only pairs that fuse, or only pairs that do not fuse.

The most widely used synthetic stimulus set is /ba/, /da/, /ga/, /pa/, /ta/, and /ka/ with identical fundamental frequency contours. As with the analogous natural-speech set of syllables, the reason for their popularity is primarily the convenience and availability of a stimulus set that tends to give reliable REA's—not their tendency to fuse, which has been given little attention. The differences between these stimuli are confined to the first 50 msec or so, which carry the consonantal distinctions. The vowel portions—which may last for another 250 msec or so—are exactly identical and therefore fuse perfectly in dichotic presentation. This alone is sufficient to guarantee that dichotic pairs of these stimuli will sound more or less fused (Halwes, 1969). The "more or less" will depend on the spectral similarity of the initial 50 msec. Synthetic /ba/, /da/, and /ga/, if synthesized so they differ only in the transitions of the second (and third) formant, fuse perfectly into a single syllable. This was experimentally demonstrated by requiring subjects to discriminate dichotic pairs from binaural (identical) pairs of stimuli from the same set. Most of the subjects, including experienced listeners, performed at chance level (Repp, 1976b). It is justified, therefore, to call these stimulus pairs "perfectly fused."[6]

Informal observations suggest that strong fusion is also obtained with the voiceless set (/pa/, /ta/, /ka/) if the stimuli differ only in their formant transitions, although the initial random noise (burst and aspiration) portions no longer fuse perfectly. On the other hand, stimuli that contrast in voicing (and thus in the relevant cue, voice onset time, so that a periodic waveform in one ear is accompanied by filtered noise in the other ear during the first 50 msec or so) are sufficiently different to prevent perfect fusion. The listener has some indication that different events have occurred in the two ears, but since these events are immediately followed by a perfectly fused vowel, their discrepancy is perceived only as a brief noise or roughness accompanying the perception of a single fused syllable which can be identified without great difficulty. Dichotic pairs consisting of a single phonetic percept accompanied by an auditory signal of interaural discrepancy may be called "partially fused" (cf. Repp, 1977a, 1977c).

The fusion of synthetic syllables can be effectively prevented by presenting them at different fundamental frequencies (Halwes, 1969; Repp, 1976a). Temporal asynchrony also reduces fusion, but as long as the signals overlap, they may still fuse partially. Some researchers have paired CV syllables that contrasted in their vowels as well as in the initial consonants (Studdert-Kennedy, Shankweiler, and Pisoni, 1972). Different vowels with the same fundamental frequency seem to fuse quite well, although they may be discriminable from binaural stimuli if they are spectrally dissimilar

(Kuwahara and Sakai, 1976). The frequency of the first formant may play a role in addition to fundamental frequency, but little work has been done on the fusion of complex sounds such as vowels. The influence of various other parameters (such as bursts, nasal resonance, frication, etc.) on dichotic fusion of speech sounds has not been systematically studied. If stimuli involving such differences are to be used for assessing ear advantages, their degree of fusion should first be determined.

B. The single-response paradigm

The standard procedure requires the subjects in a dichotic test to identify both competing stimuli. While appropriate with unfused stimuli, the two-response procedure has also been used with synthetic syllables subject to dichotic fusion (e.g., Shankweiler and Studdert-Kennedy, 1975). It is not surprising that the overall accuracy was quite low in these studies, because at least one of the two responses must have been a guess. Although it is possible to analyze only first responses and ignore second responses, one cannot be sure that the subjects always record their most confident response first, even when instructed to do so. Thus, the responses reflecting what the listeners actually perceived are distributed over two response columns, and it is impossible for the experimenter to identify them reliably. Hence, instructions to identify two stimuli when only one is heard are inappropriate. The only appropriate instruction is simply to identify the syllable heard (the single-response paradigm). The listener need not even be informed about the presence of different events in the two ears. Instructions to selectively attend to one ear are also inappropriate when the stimuli are fused, since it has been shown that selective attention to one ear has little or no effect with fused stimuli (Halwes, 1969; Repp, 1976b). The topic of selective attention will be discussed in more detail in Sec. IV.

Thus, dichotic tests using fused syllables are quite different from those using unfused stimuli. With unfused stimuli, the subject gives two responses which are then classified as correct or incorrect. The emphasis is on accuracy of identification. A large number of errors is desirable. These errors should be due to dichotic competition only; the monaural intelligibility of the stimuli should be as high as possible. The "raw" ear advantage (d) is defined as the difference between the proportions of correct responses for the two ears.

In a test using fused stimuli, on the other hand, only a single response is given to each stimulus pair. Ideally, this response should match one or the other of the component stimuli. Dichotic pairs for which this indeed tends to be the case (e.g., /ba/–/da/, which is heard as either /ba/ or /da/) are especially desirable. Other pairs also yield hybrid responses such as "psychoacoustic fusions" or blend responses (Cutting, 1976; Repp, 1976b, 1977a). The methodological problems created by such responses will be discussed in Sec. IV. If we consider only the "ideal" pairs, such as /ba/–/da/, where virtually all responses match one of the two component stimuli, we see that there are no errors and accuracy is perfect (or, in practice, as good as the monaural intelligibility of the stimuli). The question is not how accurately each ear performed but how the competing information was weighted and combined into a single perceptual outcome. Thus, the emphasis is on dichotic integration, not on competition. Instead of different accuracy levels for each ear, we have two complementary proportions representing each ear's share of the responses. The difference between these proportions represents the "raw" ear advantage.

Despite the theoretical and methodological differences, the two paradigms also have much in common. Specifically, the problems encountered in deriving an appropriate laterality index are rather similar. This will become evident in the following section which derives such an index for the single-response paradigm.

III. LATERALITY INDICES FOR THE SINGLE-RESPONSE PARADIGM

A. Ear dominance and stimulus dominance

In this section, the simplifying assumption is made that each stimulus pair in a test using fused stimuli yields only two kinds of (single) responses, one that matches the stimulus presented to the left ear and one that matches the stimulus in the right ear. One example, already mentioned in the preceding section, is the pair /ba/–/da/, which is heard as either /ba/ or /da/. (For other examples, see Sec. IV E.) Thus, the responses can be divided into those reflecting perceptual dominance of the left-ear stimulus and those reflecting perceptual dominance of the right-ear stimulus. Taking into account the two possible channel/ear assignments of the stimuli, the data for a single stimulus pair can then be represented in a 2×2 table, as illustrated in Table II. The two different channel/ear assignments of the stimuli constitute the rows of this table, and the two responses the columns. The entries are the proportions of the two responses for each of the two channel configurations.

Perceptual dominance is a probabilistic phenomenon, so that, in general, both responses will occur with some frequency over a number of single-response trials. There are two independent factors that determine which of the two competing stimuli dominates the perception of the fused syllable at a given time. One is the tendency of (the stimulus in) one ear to dominate (the stimulus in) the other ear. It is appropriately called *ear dominance* and, of course, is analogous to the ear advantage observed in the two-response paradigm.[7] The other factor is the tendency of one stimulus to dominate the other stimulus, regardless of their particular channel assign-

TABLE II. The data structure for a single stimulus pair in the single-response paradigm, with sample values.

Channels LE – RE	Responses /ba/	Responses /da/	
/ba/ – /da/	$x_i = 0.276$	$1-x_i = 0.724$	1.000
/da/ – /ba/	$y_i = 0.487$	$1-y_i = 0.513$	1.000

ment. It has been termed *stimulus dominance* and constitutes an important phenomenon in its own right (Repp, 1976b, 1977a, 1977c).

The two factors are illustrated by the fictitious data in Table II. Ear dominance is reflected in the difference between the averages of the diagonal entries in the 2×2 table. In the present example, there is a right-ear dominance: $\frac{1}{2}(72.4+48.7)=60.5\%$ of the responses went to the right ear and only $\frac{1}{2}(27.6+51.3)=39.5\%$ to the left ear. At the same time, there is a pronounced stimulus dominance effect, which is reflected in the difference between the column averages: /da/ was heard in $\frac{1}{2}(72.4+51.3)=61.8\%$ of the trials; /ba/ only in $\frac{1}{2}(27.6+48.7)=38.2\%$.[8]

It should be emphasized that the information about ear and stimulus dominance is contained only in the complete 2×2 contingency table but not in its individual rows. The two different channel assignments of a particular stimulus pair must always be considered together; otherwise, the results can be very misleading. In Table II, for example, /da/–/ba/ (with /ba/ in the right ear) shows a slight LEA, while /ba/–/da/ (with /da/ in the right ear) shows a very large REA. Such an apparent reversal of the ear advantage can appear puzzling if it is interpreted without considering the joint operation of two factors, ear dominance and stimulus dominance (cf. Speaks, Niccum, Carney, and Marble, 1975; Niccum, Speaks, and Carney, 1976). In fact, the right-ear dominance underlying these data is canceled by stimulus dominance in the pair /da/–/ba/, and it is augmented by stimulus dominance in the pair /ba/–/da/. Neither case in isolation reveals the actual size of the REA which lies between these extremes and must be inferred from the complete contingency table. Likewise, an appropriate estimate of stimulus dominance in an individual stimulus combination can only be derived from the complete table.

Table II bears a close resemblance to the 2×2 contingency table for the two-response paradigm (Table I). However, in Table I, the dimensions were left/right ear and correct/incorrect responses. The analogy becomes closer if we arbitrarily consider one response in Table II as "correct" (e.g., /ba/) and the other as "incorrect" (e.g., /da/). The rows of the two tables remain incompatible, however: In Table I, they represent the individual component stimuli in each ear, while, in Table II, they represent the two possible channel/ear assignments of both component stimuli. In the two-response paradigm, it is easy to summarize the responses to all stimulus pairs in a single table; in fact, it is standard procedure to do so, and the data are rarely broken down to the level of individual stimulus pairs. Basically, each individual channel assignment of each stimulus pair yields its own 2×2 table (of the form shown in Table I), and these tables are then simply added up or averaged. This presents no problem, because each stimulus pair yields left-ear and right-ear as well as correct and incorrect responses. In the single-response paradigm, on the other hand, the 2×2 tables for the individual stimulus pairs are not commensurate—their rows and columns have different labels in each case— and therefore cannot simply be added up or averaged. Even if we stipulate that the positive diagonal always contain right-ear responses and the negative diagonal left-ear responses (as in Table II), there remains one degree of freedom for the arrangement of the table. The next section shows how this problem can be solved.

B. The *e* index for the single-response paradigm

The problem facing us is how to compute an appropriate laterality index for a whole single-response test. It is easy to compute ear dominance indices for individual stimulus pairs. Despite the different nature of the entries in Table I and Table II, the structure of the data is almost completely identical in the two cases, and most of the discussion of Sec. I applies. In particular, the factor of stimulus dominance exacts the same constraints here as the factor of performance level in the two-response paradigm. A 50/50 distribution of responses here is analogous to a 50% performance level there. The simple difference index, $d_i=y_i-x_i$, is unsatisfactory for the same reason that d is unsatisfactory in the two-response paradigm. (The subscript i indicates that we are dealing with a single stimulus pair.) Clearly, the best choice is

$$e_i=(y_i-x_i)/(y_i+x_i) \quad \text{if } \tfrac{1}{2}(y_i+x_i)\le 0.5\ ,$$
$$e_i=(y_i-x_i)/(2-y_i-x_i) \quad \text{if } \tfrac{1}{2}(y_i+x_i)\ge 0.5\ . \qquad (21)$$

Since the arrangement of the 2×2 data table is arbitrary, we may adopt the convention of tabulating the less frequent response in the left column (as in Table II), so that the first condition always holds and

$$e_i=(y_i-x_i)/(y_i+x_i)\ . \qquad (22)$$

Thus, a laterality index can be computed for each individual stimulus pair. The most straightforward way of arriving at an index for the whole test would then be to take the average of all the e_i indices. However, these indices vary considerably in their precision, depending on how much stimulus dominance deviates from equilibrium. e_i indices are most reliable when the two stimuli are in equilibrium, and they become more variable and unreliable as the relative dominance of one or the other stimulus increases (cf. Sec. I C). This follows straightforwardly from statistical arguments. Therefore, e_i indices for stimulus pairs with very asymmetrical response distributions should receive less weight than indices for stimulus pairs with more nearly symmetrical response distributions. The degree of asymmetry is represented by the proportion of the less frequent of the two responses, $w_i=\frac{1}{2}(y_i+x_i)$, which is the appropriate weight to be assigned to each e_i. We then compute the overall E index as the weighted average of the e_i indices,

$$E=\sum w_ie_i\Big/\sum w_i$$
$$=\frac{1}{2}\sum\left[(y_i+x_i)(y_i-x_i)/(y_i+x_i)\right]\Big/\frac{1}{2}\sum(y_i+x_i) \qquad (23)$$
$$=\sum(y_i-x_i)\Big/\sum(y_i+x_i)=e\ .$$

Thus, the result turns out to be identical with the e index computed from a summary 2×2 table for the whole test. Note that, by adopting the conventions of tabulating the less frequency response in the left column *and* right-ear responses in the positive diagonal, we have fixed the format of the data tables, so that they can now be added up or averaged in a nonarbitrary way. The e index computed from this summary table is then identical with the weighted average of the e_i indices for the individual stimulus pairs.

The variance of the e_i indices provides us with an estimate of whether the overall e index is significantly different from zero. Assuming that the e_i indices are approximately normally distributed around zero if the null hypothesis is true, we make use of the well-known relation that the estimated variance of the mean is the sample variance divided by the number of observations,

$$s^2(e) = s_w^2(e_i)/N \,, \tag{24}$$

where N is the number of stimulus pairs. The subscript w indicates that, again, we would like to assign more weight to the deviations of the more reliable indices from the mean than to the deviations of unreliable indices. We thus compute the weighted variance of the e_i indices as

$$\begin{aligned} s_w^2(e_i) &= \sum w_i(e_i - e)^2 \Big/ \sum w_i \\ &= \sum w_i e_i^2 \Big/ \sum w_i - e^2 \\ &= \sum \left[(y_i - x_i)^2/(y_i + x_i)\right] \Big/ \sum (y_i + x_i) \\ &\quad - \left[\sum (y_i - x_i) \Big/ \sum (y_i + x_i)\right]^2 \,. \end{aligned} \tag{25}$$

Confidence limits for e can then be estimated by $e \pm 2s(e)$. If they do not include zero, e is significant at approximately $p < 0.05$.

C. The e' index

e will be a useful index as long as the distribution of the e_i indices is roughly symmetrical. With very asymmetric distributions, however, an arithmetic mean is not the optimal measure. There is an alternative method available which also permits an approximate graphical determination of the laterality index. This method makes use of the basic concepts of signal detection theory referred to in Sec. I. In addition, it provides a direct test of the assumptions underlying the e index, which is of great significance. The procedure is illustrated in Fig. 4 using some actual data from a recent experiment by Repp (1977a).

We again restrict our attention to the less frequent responses only, i.e., to the left columns of the N data tables for the individual stimulus combinations. The entry in the top row (x_i) represents left-ear responses, or "false alarms." The entry in the bottom row (y_i) represents right-ear responses, or "hits." We then plot y_i—or p(H), the hit probability—against x_i—or p(FA), the false alarm probability—for all stimulus pairs. This results in a swarm of points located on or below the negative diagonal of the unit square. (Therefore, only its lower triangular portion is shown in Fig. 4.) To these points, a receiver operating characteristic (ROC) function may be fitted. The standard ROC function is curvilinear, but for our purposes little accuracy is lost by simply fitting a linear function. A straight line through the origin and the data points may be fitted by eye, or, more precisely, by the method of least squares. The slope b of this line will range from infinity (perfect REA) to 0 (perfect LEA). It order to convert this range to the standard scale from $+1$ to -1, we define

$$e' = (b-1)/(b+1). \tag{26}$$

This value can also be read off a linear scale on the negative diagonal, as illustrated in Fig. 4. The triangles are the average results of eight subjects, while the circles are for a single experienced listener (the author) who showed an especially large REA. Based on 24 data points (stimulus pairs) in each case, the e' coefficients are $+0.55$ and $+0.96$, respectively. It can be seen in Fig. 4 that the point swarms are fitted reasonably well by straight lines (isolaterality contours) through the origin, which provides empirical support for the e' index. Further support has been obtained in recent experiments using dichotic voicing contrasts (Repp, 1977c). e' may be directly calculated as

$$e' = \tan\left\{\tfrac{1}{2}\arctan\left[\left(\sum y_i^2 - \sum x_i^2\right) \Big/ 2\sum x_i y_i\right]\right\} \tag{27}$$

which effectively is a rotation of the best-fitting line into the ±45 degrees sector, so that its slope (the tangent) ranges from $+1$ to -1.

e' is an unbiased measure in terms of signal detection theory, since it is a simple linear transformation of the area under the ROC function, a commonly used measure of sensitivity that is independent of any particular assumptions about the internal representations of the sensory events (Green and Swets, 1966; Richardson, 1972). It is monotonically related to the signal-detection index d'. Testing the significance of e' is not straightforward, however, so that one may rely on the e approximation [Eq. (22)] for this purpose.

e' is usually well approximated by e. Table III presents e' and e coefficients, together with $s(e)$, for the eight subjects in Repp's (1977a) study. It can be seen that e is generally very close to e'; the largest deviation

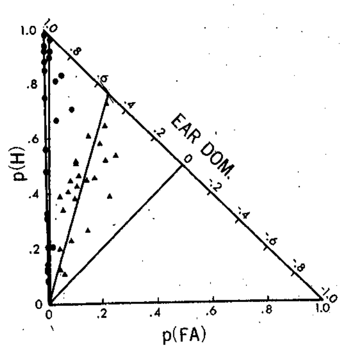

FIG. 4. Illustration of the graphic derivation of the e' index. Data from Repp (1977a). See text for explanation.

TABLE III. Ear advantages on the voicing dimension. Data of eight subjects from Repp (1977a).

Subjects	e'	e	$s(e)$
JK	+0.17	+0.16	0.06
JL	+0.73	+0.72	0.06
RG	+0.89	+0.89	0.02
MR	+0.57	+0.49	0.10
GG	−0.09	−0.12	0.08
WT	+0.90	+0.88	0.04
TJ	+0.47	+0.44	0.07
CW	+0.75	+0.74	0.06

occurs for subject MR, who in fact showed a highly asymmetrical distribution of e_i indices. By the $\pm 2s$ criterion, all coefficients except that for subject GG (the only case of left-ear dominance) are significant. It should be noted that $s(e)$ becomes constrained as e approaches ± 1 (cf. subjects RG and WT in Table III), so that it should not be used for testing whether two coefficients are significantly different from each other. A nonparametric test may be used for this purpose.

One important difference between the present procedure of deriving e' and the signal detection paradigm should be pointed out. In the latter, "bias" is varied by means of instructions, payoffs, etc., while the stimuli for which sensitivity is being measured are held constant. If the stimuli (e.g., signal and/or noise levels) were to be changed, the listener's sensitivity would change, too. In the present case, stimulus dominance takes the role of bias and ear dominance that of sensitivity. However, in order to change stimulus dominance, the stimuli themselves are varied. Thus, it is assumed that ear dominance is independent of the nature of the stimuli, at least within a given class (such as initial stop consonants). The validity of this assumption is an empirical question. It is especially convenient that determining e' for a set of data at the same time provides a test of its underlying assumptions. So far the results have been encouraging. Moreover, no correction for guessing is needed for e' since, in general, random guessing plays only a small role in the single-response paradigm. However, the single-response paradigm is not without its own problems. The last section provides a discussion of a number of methodological issues and problems not considered so far.

IV. PROBLEMS IN MEASURING THE DICHOTIC EAR ADVANTAGE

A. Stimulus intelligibility

It is good practice to precede a dichotic test with a series of binaural (or monaural) stimuli, in order to familiarize the listener with their sound and to find out whether they can be reliably identified. In order to obtain useful dichotic data, the stimuli must be intelligible and yield high binaural (or monaural) identification scores.[9]

This goal is more easily achieved with natural speech stimuli than with synthetic speech. However, synthetic stimuli are desirable because their acoustic properties can be controlled by the experimenter. Therefore, it is advisable to use a good set of synthetic stimuli that has been pretested for intelligibility—a point that has often been neglected in the past.

Even when the average intelligibility of a set of syllables is high, their intelligibility should be tested for each individual subject in a given test. From time to time, individuals are encountered who find it very difficult to identify synthetic speech sounds. Such individuals may have to be excused from the test. (This is an obvious problem in clincial applications of dichotic tests using synthetic syllables.)

Intelligibility is usually assessed in terms of the confusions that occur between members of a stimulus set. The information obtained from a binaural (monaural) confusion matrix may be used to apply a correction to dichotic data which leads to a better estimate of the stimulus dominance relationships between the stimuli (Repp, 1976b). Unfortunately, however, information about ear dominance cannot be recovered in this fashion—confusable stimuli yield smaller ear advantages than nonconfusable stimuli (Repp, 1977a). Since this effect may be confounded with individual differences in confusion patterns, it is advisable to omit confusable pairs when calculating ear advantage indices for individual subjects.

Problems arising from confusability of certain stimuli may also be reduced by using a dichotic listening procedure that does not require a labeling response. I mention here only one such alternative, originally proposed by Preston, Yeni-Komshian, and Benson (1968), that is especially suited for fused stimuli: The two component stimuli are presented binaurally, followed by the dichotic pair, and the listener judges whether the dichotic stimulus was more similar to the first or the second binaural stimulus. Repp (1977c, and work in progress) has used an AXB version of this ABX paradigm, i.e., with the dichotic pair in the middle of each stimulus triad. This method may yield cleaner data than the single-response identification task, but it is more time consuming.

The intelligibility–confusability issue raises an important theoretical problem. Individual differences in the perception of stimuli (especially of synthetic syllables) are large, and "poor subjects" who produce many confusions will tend to have smaller ear advantages than "good subjects." The individual differences that thus confound the measure of the ear advantage may be due to differential sensitivity to the acoustic cues manipulated in synthetic speech, which are just a subset of the cues available in natural speech for the relevant phonetic distinctions. The cues of synthetic stimuli may be perfectly sufficient for some listeners, but only minimally so for others. Now suppose we have succeeded in generating an excellent stimulus set that produces no confusions at all. Do we have eliminated the individual differences? Overtly, yes; but if the stimuli were attenuated or mixed with white noise, some subjects probably would produce more confusions than others. Also, if tested with an acoustic stimulus continuum as used in categorical-perception studies, some subjects would have sharper category boundaries than others. These individual differences would most likely reflect the same differences in cue sensitivity that are evident with inherently confusable stimuli.

It seems quite possible that such individual differences in perceptual variability play a role in the perception of dichotic stimuli. A fused dichotic syllable pair is often quite ambiguous, and an unfused stimulus pair is often degraded through mutual interference between the two stimuli. The problem is best illustrated with a fused pair, e.g., /da/–/ga/, as shown in Fig. 5. Assuming perfect intelligibility of the component stimuli and no pronounced stimulus dominance effect, this dichotic pair sometimes sounds like /da/ and sometimes like /ga/. (Because of categorical perception, the subject may often not be aware of the inherent ambiguity of the syllable.) For individuals with a REA, the dichotic pair sounds a little more (often) like /da/ when /da/ is in the right ear, and a little more (often) like /ga/ when /ga/ is in the right ear. Thus, the two fused stimuli may be considered as lying on a /da/–/ga/ continuum, a little to the left and a little to the right of the category boundary, respectively. A "good subject" with low perceptual variability has a sharp category boundary and thus resolves the two dichotic stimuli well; she will show a clear REA. A "poor subject", on the other hand, *with the same degree of lateralization* as the good subject, is likely to have a flatter psychometric function separating the two categories and, as a result, will produce similar response distributions for the two dichotic pairs and thus a smaller REA. This is schematically illustrated in Fig. 5.

If this argument is correct, it implies that individual differences in the dichotic ear advantage may be inextricably confounded with individual differences in perceptual accuracy. This would be a serious obstacle to measuring individual ear advantages on an ordinal scale.

This problem seems to be less acute in the two-response paradigm; there, variations in perceptual accuracy are translated, at least in part, into variations in performance level which can be dealt with more easily. However, this apparent advantage of the two-response paradigm is offset by a number of disadvantages that are discussed in the next paragraphs.

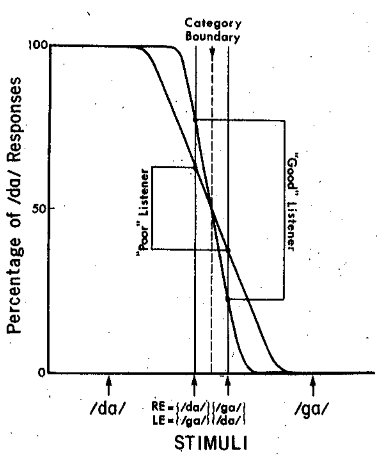

FIG. 5. Schematic illustration of the effect that individual differences in perceptual accuracy might have on ear dominance for fused syllables.

B. Stimulus dominance

We have seen that the factor of stimulus dominance can be dealt with elegantly in the single-response paradigm. Stimulus dominance in this paradigm is somewhat analogous to performance level in the two-response paradigm. However, this analogy is purely formal—i.e., these are the factors that can be effectively handled in the respective paradigms by using similar methods—, but they are conceptually very different. As the preceding paragraphs have shown, there are presumably variations in "performance level" (i.e., perceptual accuracy) in the single-response paradigm, but they are covert and much more difficult to deal with. Conversely, there is the problem of how to deal with stimulus dominance in the two-response paradigm.

Although stimulus dominance may be expected to play a smaller role in the two-response paradigm, there is evidence that it is nevertheless present. Berlin *et al.* (1973), for example, have reported that unfused natural-speech syllables contrasting in voicing receive more correct voiceless responses than correct voiced responses. Berlin *et al.* reduced this asymmetry by aligning the stimuli at the first pitch pulse rather than at stimulus onset. Speaks *et al.* (1975), who used the same alignment criterion, reported data suggesting that stimulus dominance effects are relatively small with natural-speech stimuli. However, this issue needs to be investigated in more detail. There is no doubt that strong stimulus dominance reduces the manifest ear advantage, and if some individuals show stronger effects than others, these individual differences are confounded with the measure of the ear advantage. At present, there is no way of dealing with this potential problem.

C. Guessing

Random guessing plays an insignificant role in the single-response paradigm. Random guesses following lapses of attention or highly ambiguous stimuli may occur now and then, but, in general, the listener reports what she hears and does not resort to guessing (except for "sophisticated" guessing between a very limited number of alternatives). In the two-response paradigm, on the other hand, guessing is commonplace. Frequently a listener can identify the stimulus in one ear but has few or no clues about the stimulus in the other ear. The resulting guesses cannot be reliably identified in the data and lead to a considerable amount of random variation. The correction for guessing proposed in Sec. I D is a rather crude procedure (cf. also Repp, 1977b), and alternative ways of dealing with the guessing problem should be considered.

One obvious possibility is to instruct the listeners not to guess, i.e., to give zero, one, or two responses per stimulus pair, depending on how many stimuli can be identified with reasonable confidence. This method has rarely been used, because the resulting heterogeneous protocols are difficult to analyze. More commonly, subjects are instructed to write down the more confident response first, and then only these responses are analyzed. Effectively, this is the single-response paradigm applied to unfused stimuli. (The second response might just as well be omitted.) If both stimuli can be identified on a trial, the procedure amounts to a judgment which of them is the more salient. This procedure is interesting because it reduces guessing and permits the methods of Sec. III to be applied, so that stimulus dominance can be taken into account. The main problem is the control of selective attention, discussed in the next paragraphs.

D. Selective attention

The most important difference between fused and unfused syllables lies in the effect of selective attention. Perfectly fused syllables are heard as originating in the middle of the head, and voluntary efforts to pay attention to one ear has no effect on the responses (Repp, 1976b). The effect of selective attention with partially fused syllables appears to be very small (see Halwes, 1969; Repp, 1976a).[10] Unfused syllables, on the other hand, yield large attentional effects, and practiced listeners are able to reach almost perfect scores when reporting only the syllables in one ear (Halwes, 1969). It is fair to say that the effectiveness of selective attention is a direct function of the degree of fusion of two stimuli (cf. also footnote 4).

It follows that, with unfused syllables, it is not possible to separate attentional preferences, effectiveness, or bias from the ear advantage per se, which presumably has a physiological basis. Some researchers have hypothesized that the ear advantage is entirely an attentional phenomenon (Kinsbourne, 1973, 1975; Morais and Landercy, 1977) or a perceptual advantage for stimuli localized to the right of the midline (Morais and Bertelson, 1973, 1975; Morais, 1975; Hublet, Morais, and Bertelson, 1976). The fact that a significant REA is obtained for perfectly fused syllables in the absence of any attentional effects (Repp, 1976b, 1976c) suggests that there are both physiological and attentional components of the ear advantage. It is possible that perfectly fused syllables yield an estimate of the physiological component alone, with the attentional component removed. With unfused syllables, physiological and attentional effects are confounded.

In fairness, one should distinguish two kinds of attentional effects: automatic and strategic biases. Automatic biases may arise from contingencies and expectancies within the experimental situation; for example, during or after processing a verbal stimulus, the left hemisphere may be activated more than the right, which leads to an automatic bias for stimuli on the right side. These kinds of involuntary biases are what Kinsbourne and Morais have in mind. The REA for unfused syllables apparently can be influenced by contextual factors (Goldstein and Lackner, 1974; Morais and Landercy, 1977); whether the same is true with fused syllables remains to be investigated. Individual differences in automatic attentional preferences are difficult to distinguish from differences in the physiological or functional ear asymmetry itself; they may be highly correlated. Strategic biases, on the other hand, are voluntary and at the disposition of the listener. For example, by deliberately paying attention to the left ear, even persons with a strong REA can produce a LEA with unfused stimuli. Such strategies are not under control in the standard two-response paradigm, so that the ear advantages obtained are not a pure measure of lateral asymmetry.

This is especially obvious when the single-response paradigm is applied to unfused stimuli. It does not suffice to instruct the listeners to pay equal attention to both ears; unexperienced listeners may not follow these instructions, and there is no way of controlling whether they do. It may be difficult in principle to "neutralize" attention. Requiring two responses at least reduces the effect that attentional biases would have in the single-response paradigm. There remains the possibility of controlling the listener's strategies by instructions to pay attention to one ear and to report only the stimuli in that ear. This procedure has been followed by several researchers, although usually not for the purpose of assessing ear advantages. It may be considered as a two-response paradigm in two passes; in this case, a single ear advantage index would be computed after combining the results of two (properly counter-balanced) selective-attention conditions. The problem is here that, because of the relative efficiency of selective attention, performance level will be rather high, making the ear advantage index less reliable. Alternatively, the two selective-attention conditions may be considered as single-response paradigms, and two separate single-response indices may be computed whose difference is then taken as the measure of the ear advantage. However, here we encounter the same problem as with the *d* index: Simple differences depend on the absolute size of the numbers involved, so that the resulting index would reflect individual differences in the relative effectiveness of selective attention in addition to the ear advantage itself. Regardless of the form of data analysis, there is the theoretical possibility that there are lateral asymmetries in the effectiveness of voluntary selective attention which are independent of the ear advantage itself and again would confound the measure of the ear advantage.

It must be concluded that there is no perfect way of controlling attentional strategies with unfused stimuli, so that fused stimuli offer a significant methodological advantage in this respect.

E. Blend responses

In the discussion of the single-response paradigm (Sec. III), it has been assumed that only two kinds of responses are given to a fused dichotic pair; they match one of the two component stimuli and can be assigned to one or the other ear. Of the fifteen possible combinations of the six standard stop-consonant-vowel syllables, only seven meet this strict criterion, given that they are highly intelligible in isolation. These pairs are the place contrasts /ba/–/da/, /da/–/ga/, /pa/–/ta/, and /ta/–/ka/, and the voicing contrasts /ba/–/pa/, /da/–/ta/, and /ga/–/ka/. These are the stimulus pairs especially suited for the methodology outlined in Sec. III.

However, it may be desirable for some purpose to include other stimulus combinations as well, and past experiments have almost always included all possible combinations of the stimuli. The two place contrasts, /ba/–/ga/ and /pa/–/ka/, may receive a third response, /da/ and /ta/, respectively. Cutting (1976) has called these intermediate percepts "psychoacoustic fusions." Their frequency may be negligible for many listeners, but some subjects give a substantial proportion of these responses (Repp, 1976b). The remaining stimulus combinations are the six double-feature contrasts: /ba/–/ta/, /ba/–/ka/, /da/–/pa/, /da/–/ka/, /ga/–/pa/, and /ga/–/ta/. They typically yield two additional responses per pair, resulting from the combination of the feature values of the component stimuli; e.g., /ba/–/ta/ is heard not only as /ba/ or /ta/ but also as /pa/ and /da/. These "blend" responses are usually quite frequent and may even exceed the proportions of correct responses, although there is much variation between stimulus pairs and subjects in this respect (Halwes, 1969; Repp, 1977a). Blend responses and psychoacoustic fusions do not convey direct information about ear asymmetries, so that the question arises what to do with them.

Hybrid responses also occur with unfused syllables (Halwes, 1969; Studdert-Kennedy and Shankweiler, 1970). In the two-response paradigm, they are simply grouped together with other types of errors in the class of incorrect responses. As a result, double-feature contrasts typically have higher error rates than single-feature contrasts, an effect that has been termed the "feature-sharing advantage" (Studdert-Kennedy and Shankweiler, 1970; Studdert-Kennedy *et al.*, 1972; Pisoni, 1975). The availability of the correct-incorrect distinction makes it easy to dispose of blends in the two-response paradigm.

In the single-response paradigm, on the other hand, we have assumed that all responses are "correct," apart perhaps from a few random errors (which may be divided up randomly between the two response categories). Blend responses are different from random errors in that they reflect what the listener actually heard; in a sense, they are correct responses. However, they cannot be unambiguously assigned to one or the other ear. There are two ways of dealing with them. One possibility, followed by Repp (1977a), is to analyze the data in terms of the individual phonetic features and to calculate two laterality indices, one for voicing and one for place. If only one feature is considered at a time, blend responses become informative with respect to lateral asymmetries. The two resulting indices may be averaged to obtain a single index.

The other possibility, which consists in discarding blend responses, is more problematic. Consider the following extreme example, shown in Table IV. Here omission of blends (/pa/ and /da/ responses) would

TABLE IV. Fictitious response distribution for a double-feature contrast pair.

Stimuli LE – RE	Responses /ba/	/ta/	/pa/	/da/
/ba/–/ta/	0	0.33	0.33	0.33
/ta/–/ba/	0.5	0	0.25	0.25

lead to the conclusion that there is a perfect REA for this stimulus pair. However, if the data are analyzed for each feature separately, it is found that there are only moderate REA's for voicing and place ($e = +0.46$ for both). If blend errors are discarded, this information is lost and the REA is inflated. It is not clear which should be considered the correct index: the average of the separate indices for the two features or the index based on "correct" responses only.

It may be possible to settle the problem by examining empirical isolaterality contours (ROC functions) for single- and double-feature contrast pairs. In the meantime, double-feature contrasts and pairs yielding psychoacoustic fusions are best omitted from dichotic single-response tests, as long as only the ear advantage is of concern. This leaves us with only seven of the original fifteen stimulus pairs—perhaps too few to constitute a useful test. However, synthetic stimuli offer the possibility of varying the acoustic structure of the stimuli while leaving their phonetic content unchanged. By varying voice onset time or the formant transitions within phonetic categories, stimulus dominance relationships can be changed (Repp, 1976b, 1977a, 1977c). It is possible to take a single stimulus pair (e.g., /ba/–/pa/), to select several tokens with different acoustic characteristics (e.g., several different voice onset times within each category), to combine them dichotically, and thus to arrive at a test that contains a sufficient number of stimulus pairs varying in stimulus dominance relationships, is maximally homogeneous, and leads to a clean estimate of ear dominance (Repp, 1977c).[11] This illustrates one of the great methodological advantages of the single-response paradigm over the two-response paradigm; in the former, two response alternatives are perfectly sufficient, while the latter always requires a larger number of response alternatives in order to reduce the effect of guessing.

F. Test reliability

In the Introduction, it has been stressed that dichotic tests must satisfy general test-theoretical standards. One of these is reliability. As in any other psychological test, the observed score (ear advantage) of a subject represents the "true" score plus random measurement error. The magnitude of the measurement error depends on the length of the test. It is not surprising that, in repeated administrations of a short dichotic test, the observed ear advantages for a given subject vary considerably and may even show reversals in direction (Speaks, Niccum, and Carney, 1976). Most dichotic studies in the past have used short tests whose reliability was likely to be low. The fact that a certain percentage of right-handed subjects show either no REA or a LEA (although physiological data suggest that virtually all are left-hemisphere dominant for speech) is at least in part due to measurement error (cf. Blumstein, Goodglass, and Tartter, 1975).

Ryan and McNeil (1974) reported a test-retest reliability coefficient of +0.80 for a 60-item test; Blumstein *et al.* (1975) found a somewhat lower coefficient of +0.74 for an 80-item test; Catlin, VanDerveer, and Teicher (1976) found an even lower reliability of +0.64 for a 120-item test.[12] The first two studies used natural-speech CV syllables and the third (probably fused) synthetic syllables in the two-response format. These reliabilities are quite satisfactory in view of the relative shortness of the tests and the weaknesses of the two-response paradigm (guessing, attentional fluctuations, etc.). Researchers in the field have tended to expect too much from a short dichotic test and have been reluctant to accept the conclusion that much longer tests will be necessary to obtain precise measurements. If we accept the Blumstein *et al.* results as typical and apply the standard Spearman–Brown formula (Lord and Novick, 1968, p. 112), we find that the test has to be three times as long (about 240 pairs) to achieve a reliability of +0.90, and six times as long (about 480 pairs) to reach $r = +0.95$. From the Ryan and McNeil data, we obtain more moderate estimates of 140 and 280 pairs, respectively. Considering the fact that the standard set of six CV syllables yields a basic test unit of 30 dichotic pairs, it might be recommended that ten repetitions of this test unit (i.e., 300 pairs) be administered in order to obtain stable ear advantage indices. Such a test requires about 20 min of listening time and therefore should be feasible under most circumstances, both in and outside the laboratory.

Underlying the development of the single-response methodology is the hope that this procedure will prove to be more reliable than the traditional two-response paradigm. Encouraging results have been obtained recently by Bruce Wexler (personal communication, 1976). Using a 60-item test of relatively unfused syllables in a single-response paradigm, Wexler obtained reliabilities well above +0.90 for both normal and psychotic subjects. Repp (1977c) obtained reliabilities better than +0.95 with partially fused syllables (voicing contrasts); however, these tests included over 300 stimulus pairs. These results were obtained with very small subject samples, and further research will be required to properly assess the reliability of these methods. Alternative methods, such as the AXB paradigm mentioned earlier, may also lead to increased reliability.

G. Test homogeneity and validity

The problem of test reliability is a practical one that always can be solved by using a test of sufficient length, as long as there are true individual differences. More important is the theoretical question of what is actually being measured—the validity of the test. Ultimately, its validity as an instrument for assessing hemispheric dominance needs to be assessed by physiological criteria of functional lateralization. At present, however, these physiological measurements are still crude and hazardous; moreover, they are a less crucial criterion than they may seem at first thought. First of all, the only reliable physiological indicator in normal subjects, the Wada test, yields only categorical outcomes (left, right, or no dominance), not a graded scale of lateralization. Moreover, it really supplies a useful criterion only for the small group of left-handers, since it is now well established that virtually all right handers are left-hemisphere dominant for speech. Secondly, the original idea that the dichotic ear advantage directly reflects hemispheric dominance for speech is probably an oversimplification. It is likely that there are multiple factors underlying the dichotic ear advantage, only one of which is the (quite possibly all-or-none) dominance of one hemisphere for speech. The primary task of the theoretical study of the ear advantage must therefore be to determine what is actually being measured. This is a difficult problem, but some preliminary steps are possible by asking the following familiar test-theoretical questions: Do all items in a test measure the same underlying variable(s)? Do different tests composed of different types of items from the same general class (viz., those that tend to yield an average REA) measure the same underlying variable(s)? Do different paradigms measure the same underlying variable(s)?

These important (and closely related) questions about within-test (or item) homogeneity and between-test homogeneity (or validity) have been given little thought in the past. Their answers are by no means obvious. Consider the question of item homogeneity. Repp (1976b), for example, found that two fused stimulus pairs of a three-item test yielded REA's but the third pair did not. The cause for this is not known, and more evidence is needed. The statistical techniques that may be applied are intercorrelation of laterality indices for individual items in a test and subsequent factor analysis, or perhaps some adaptation of the more recent methods of stochastic test theory (Rasch, 1960; Lord and Novick, 1968). These analyses should determine whether all items in a test measure the same factor, or whether different items measure different factors. The empirical derivation of an isolaterality function in the single-response paradigm (Sec. III B) also constitutes a (less rigorous) test of item homogeneity. (However, even if it turned out that only a single factor is being measured, this would show only that the test is homogeneous and all items measure the same thing; the single factor may nevertheless represent a complex of underlying variables.)

A related problem is whether the ear advantages for different phonetic features reflect the same underlying factors. In a recent study of partially fused dichotic double-feature contrasts, Repp (1977a) calculated e coefficients separately for voicing and place; they correlated only +0.64, although each index was based on the same 768 trials. Repp hypothesized that individual differences in perceptual organization may be reflected in the dichotic ear advantage (see also Sec. IV A). Tests of this hypothesis are needed.

Finally, the homogeneity question needs to be asked about whole tests: Do tests composed of different types of speech stimuli (e.g., CV's, VC's, VCV's, or words; stops, fricatives, or nasals; etc.) measure the same factor? Do tests composed of natural-speech syllables measure the same factor as synthetic tests? Do fused and unfused syllables (or, the single-response and the two-response paradigm) assess the same factor? Again, intercorrelations between different tests (perhaps supplemented by factor analysis and modern test theory) should provide an answer. So far, there are no data available. A positive result would reassure us that we are actually measuring a well-defined characteristic whose complexities will have to be unraveled primarily by physiologists. Negative results, on the other hands, disastrous as they would be for the diagnostic application of dichotic tests, would be of great theoretical interest. Perhaps there is more than one "ear advantage," i.e., different tests may tap different dimensions of a very complex phenomenon (cf. Porter and Berlin, 1975).

H. Absolute magnitude of the ear advantage

The questions of reliability, homogeneity, and validity, which are correlational in nature, must be kept separate from the issue of the absolute magnitude of the ear advantage. For example, ear advantages may increase with practice (although the evidence appears to be negative—see Porter, Troendle, and Berlin, 1976), but as long as they do so for all individuals, the reliability of the test will not be affected. Different items in a test may yield different magnitudes of ear advantages, but they nevertheless may measure the same underlying variable (plus some other factor responsible for the variations in magnitude). Similarly, different classes of stimuli may yield different average magnitudes of REA and nevertheless measure the same thing. As long as all individuals tested are in basically the same rank order on each test (or on each item), the homogeneity criterion is satisfied, and it is immaterial which tests or items are selected for testing persons, as long as all persons to be compared are tested with the same tests or items. The variations in the absolute magnitude of the ear advantage represent variations in item or test "difficulty," in terms of test theory. It is a separate but nevertheless important question what causes these variations in difficulty, if they exist. On the other hand, if two items or tests yield the same average REA, this implies absolutely nothing about their intercorrelation.

One striking difference in the magnitude of ear advantages has been discovered in recent research using the single-response paradigm (Repp, 1976b, 1976c, 1977a, 1977c, and work in progress): Partially fused syllables (voicing and double-feature contrasts) yield much larger ear advantages than perfectly fused syllables (place contrasts). This result is methodologically interesting, because larger ear advantages are also likely to be more reliable. Perfect fusion seems to prevent large ear dominance effects. Research is in progress to determine whether perfectly fused, partially fused, and unfused syllables actually measure the same underlying factor. The ear advantages obtained with dichotic voicing contrasts (Repp, 1977a, 1977b) are larger than any effects reported previously for normal subjects. As far as one can tell from the small subject samples tested so far, it also seems that a larger proportion of subjects shows REA's than in previous tests, which would be in closer agreement with physiological findings. The tests seem very reliable and are easy to administer and to score. Thus, the methodology described in this paper and elsewhere holds considerable promise and may constitute a genuine improvement over existing dichotic testing methods.

ACKNOWLEDGMENTS

I would like to thank Terry Halwes for numerous discussions and comments at all stages of this project. Jack Cullen, Jerre Levy, and Dennis Klatt also provided useful comments. At the outset, this work was supported by NIH Grant DE00202 to the University of Connecticut Health Center, Farmington. Later, support came from NICHD Grant HD01994 to the Haskins Laboratories.

[1]Consider, for example, two subjects, A and B, with $P_0 = 0.6$ and 0.8, respectively. Assume that $d = 0.5$ for A and $d = 0.4$ for B. Who shows the larger ear advantage? From a comparison of d indices, the answer would be A. However, B never could have reached that index because of her higher performance level which permits a maximal d of only 0.4. There is no reason why B's better performance on the test should imply that she is less lateralized than A. In fact, once performance level is taken into account, it becomes clear that B shows the maximal d for her level of performance, while A's index is considerably below the maximal d possible at $P_0 = 0.6$ ($d_{max} = 0.8$). It therefore should be concluded that, contrary to the first superficial impression, B shows a stronger REA than A.

[2]An index very similar to POC′ but based only on single-correct responses was proposed by Studdert-Kennedy and Shankweiler (1970). This index, which is a possible alternative to the "e index" described below, will not be considered here; it is briefly discussed in Repp (1977b).

[3]Jerre Levy (personal communication 1977) has demonstrated, by a statistical proof analogous to that for the ϕ coefficient (Levy, in press), that even e is intrinsically correlated with P_0. This is an important result; however, its practical significance is not quite clear at present. The actual correlation may be of negligible magnitude, and it seems that the arguments set forth in Sec. IE are valid regardless of any intrinsic correlation between e and P_0.

[4]Richardson (1976) has criticized the arguments of Marshall *et al.* and has argued for a theory-independent ordinal measure of laterality. However, his proposal seems both too pessimistic and too restrictive to be of much practical use. It should be kept in mind that the numerical scales of the ear advantage proposed here and by Marshall *et al.* are *not* assumed to have more than ordinal properties.

[5]Nevertheless, the spectral separation of the two competing signals (particularly in terms of fundamental frequency) may affect performance. Perceptual separability may be viewed as a continuum ranging from perfect fusion to perfect separation. (See also the discussion of selective attention in Sec. IV D.)

[6]The actual syllables in this experiment were /bae/, /dae/, /gae/, but the nature of the vowel is immaterial. Perfectly fused syllables have been used in a number of other studies since (Repp, 1976c, and unpublished work).

[7]In order to avoid new acronyms, the abbreviations REA and

LEA will be maintained for the corresponding trends in ear dominance.

[8]It may be argued that stimulus dominance reflects merely response bias, i.e., a stimulus-independent tendency of listeners to give one response more often than the other. However, stimulus dominance relationships can be changed by modifying the acoustic structure of the stimuli within phonetic categories (Repp, 1976b, 1977a, 1977c), so that they are at least in part stimulus dependent. Repp (1976b, 1977a) hypothesized that stimulus dominance is completely determined by the relationship of the stimuli to the listener's perceptual category prototypes. Essentially, this is a theory of response bias. Stimulus dominance may be considered as the result of the interaction between the listener's perceptual organization and the structure of the stimuli.

[9]At the intensity levels used in dichotic testing, monaural scores for the two ears are usually similar, even in clinical cases. However, it is certainly useful to obtain audiograms for each ear of a listener prior to testing, in order to detect unilateral hearing loss. Little seems to be known about the effects of moderate hearing loss on dichotic ear asymmetries.

[10]In a test where each partially fused dichotic pair contrasts only in a single cue, such as voice onset time in dichotic voicing contrasts, the contents of one channel may be *inferred* from the localization of the unfused stimulus portion (e.g., if a noise burst is heard on the right side, the voiceless member of the pair must have been in the right ear). However, this is an indirect strategy that can be prevented by appropriate instructions (cf. Repp, 1977a, 1977c). Even with selective-attention instructions, many subjects fail to discover this strategy and perform at chance level (Miller, 1977).

[11]In principle, e' can be calculated without varying stimulus dominance. However, varying the stimuli and, with them, stimulus dominance is important in order to avoid extreme dominance asymmetries due to individual idiosyncrasies, to derive an isolaterality function, and simply to provide variety in the test.

[12]Catlin *et al.* also determined the reliability of individual differences in a monaural reaction time task that showed a significant overall REA. The reliability was basically zero, and there was no correlation between individual differences in the reaction time task and in the standard dichotic test. This rules out reaction time as a measure of ear dominance.

Berlin, C. I., Lowe-Bell, S. S., Cullen, J. K., Jr., Thompson, C. L., and Loovis, C. F. (1973). "Dichotic speech perception: An interpretation of right-ear advantage and temporal offset effects," J. Acoust. Soc. Am. 53, 699–709.

Blumstein, S., Goodglass, H., and Tartter, V. (1975). "The reliability of ear advantage in dichotic listening," Brain Lang. 2, 226–236.

Broadbent, D. E. (1955). "A note on binaural fusion," Q. J. Exp. Psychol. 7, 46–47.

Broadbent, D. E., and Ladefoged, P. (1957). "On the fusion of sounds reaching different sense organs," J. Acoust. Soc. Am. 29, 708–710.

Catlin, J., VanDerveer, N. J., and Teicher, R. D. (1976). "Monaural right-ear advantage in a target-identification task," Brain Lang. 3, 470–481.

Cullen, J. K., Jr., Thompson, C. L., Hughes, L. F., Berlin, C. I., and Samson, D. S. (1974). "The effects of varied acoustic parameters on performance in dichotic speech perception tasks," Brain Lang. 1, 307–322.

Cutting, J. E. (1976). "Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotic listening," Psychol. Rev. 83, 114–140.

Franklin, B. (1969). "The effect on consonant discrimination of combining a low-frequency passband in one ear and a high-frequency passband in the other ear," J. Aud. Res. 9, 365–378.

Goldstein, L., and Lackner, J. R. (1974). "Sideways look at dichotic listening," J. Acoust. Soc. Am. 55, S10 (A).

Green, D. M., and Swets, J. A. (1966). *Signal Detection and Psychophysics* (Wiley, New York).

Haggard, M. P. (1975). "Asymmetrical analysis of stimuli with dichotically split formant information," Speech Perception, Series 2, No. 4, 11–19 (Dept. of Psychology, The Queen's University of Belfast).

Halwes, T. G. (1969). "Effects of dichotic fusion on the perception of speech," Ph.D. dissertation, University of Minnesota (unpublished).

Harshman, R., and Krashen, S. (1972). "An 'unbiased' procedure for comparing degree of lateralization of dichotically presented stimuli," UCLA Working Papers in Phonetics 23, 3–12.

Hublet, C., Morais, J., and Bertelson, P. (1976). "Spatial constraints on focused attention: Beyond the right-side advantage," Perception 5, 3–8.

Kimura, D. (1961). "Cerebral dominance and the perception of verbal stimuli," Can. J. Psychol. 15, 166–171.

Kinsbourne, M. (1973). "The control of attention by interaction between the cerebral hemispheres," *Attention and Performance IV*, edited by S. Kornblum (Academic, New York) pp. 239–256.

Kinsbourne, M. (1975). "The mechanism of hemispheric control of the lateral gradient of attention," *Attention and Performance V*, edited by P. M. A. Rabbitt and S. Dornic (Academic, London) pp. 81–97.

Kuhn, G. M. (1973). "The phi coefficient as an index of ear differences in dichotic listening," Cortex 9, 447–457.

Kuwahara, H., and Sakai, H. (1976). "Identification and dichotic fusion of time-varying synthetic vowels" (unpublished).

Leakey, D. M., Sayers, B. McA., and Cherry, C. (1958). "Binaural fusion of low- and high-frequency sounds," J. Acoust. Soc. Am. 30, 222.

Levy, J. (1977). "The correlation of the ϕ function of the difference score with performance." Cortex (in press).

Levy, J. (1977). Personal communication.

Lord, F. M., and Novick, M. R. (1968). *Statistical Theories of Mental Test Scores* (Addison-Wesley, Reading, MA).

Marshall, J. C., Caplan, D., and Holmes, J. M. (1975). "The measure of laterality," Neuropsychologia 13, 315–321.

Miller, J. L. (1977). "Properties of feature detectors for VOT: The voiceless channel of analysis," J. Acoust. Soc. Am. 62, 641–648.

Morais, J. (1975). "The effects of ventriloquism on the right side advantage for verbal material," Cognition 3, 127–139.

Morais, J., and Bertelson, P. (1973). "Laterality effects in diotic listening," Perception 2, 107–111.

Morais, J., and Bertelson, P. (1975). "Spatial position versus ear of entry as determinant of the auditory laterality effects: A sterephonic test," J. Exp. Psychol. (HPP) 1, 253–262.

Morais, J., and Landercy, M. (1977). "Listening to speech while retaining music: What happens to the right-ear advantage?" Brain Lang. 4, 295–308.

Nearey, T. M., and Levitt, A. G. (1974). "Evidence for spectral fusion in dichotic release from upward spread of masking," Haskins Labs SR-39/40, 81–89.

Niccum, N., Speaks, C., and Carney, E. (1976). "Reversals in ear advantage with dichotic listening: Effects of alignment," J. Acoust. Soc. Am. 59, S6 (A).

Nye, P., Nearey, T., and Rand, T. (1974). "Dichotic release from masking: Further results from studies with synthetic speech stimuli," Haskins Labs SR-37/38, 123–137.

Odenthal, D. W. (1963). "Perception and neural representation of simultaneous dichotic pure tone stimuli," Acta Physiol. Pharmacol. Neerl. 12, 453–496.

Perrott, D. R., and Barry, S. H. (1969). "Binaural fusion," J. Aud. Res. 3, 263–269.

Pisoni, D. B. (1975). "Dichotic listening and processing phonetic features," *Cognitive Theory*, edited by F. Restle, R. M. Shiffrin, N. J. Castellan, H. Lindman, and D. B.

Pisoni (Erlbaum, Hillsdale, NJ) Vol. I.
Porter, R. J., Jr., and Berlin, C. I. (1975). "On interpreting developmental changes in the dichotic right-ear advantage," Brian Lang. 2, 186–200.
Porter, R. J., Jr., Troendle, R., and Berlin, C. I. (1976). "Effects of practice on the perception of dichotically presented stop-consonant-vowel syllables," J. Acoust. Soc. Am. 59, 679–682.
Preston, M. S., Yeni-Komshian, G., and Benson, P. (1968). "A dichotic ABX procedure for use in laterality studies," Haskins Labs SR-13/14, 179–180.
Rand, T. C. (1974). "Dichotic release from masking for speech," J. Acoust. Soc. Am. 55, 678–680.
Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests* (Nielson and Lydiche, Copenhagen).
Repp, B. H. (1976a). "Effects of fundamental frequency contrast on discrimination and identification of dichotic CV syllables at various temporal delays," Mem. Cognit. 4, 75–90.
Repp, B. H. (1976b). "Identification of dichotic fusions," J. Acoust, Soc. Am. 60, 456–469.
Repp, B. H. (1976c). "Discrimination of dichotic fusions," Haskins Labs SR-45/46, 123–139.
Repp, B. H. (1977a). "Dichotic competition of speech sounds: The role of acoustic stimulus structure," J. Exp. Psychol. (HPP) 3, 53-70.
Repp, B. H. (1977b). "A simple model of response selection in the dichotic two-response paradigm," Haskins Labs SR-49, 187–199.
Repp, B. H. (1977c). "Stimulus dominance and ear dominance in the perception of dichotic voicing contrasts," (unpublished).
Richardson, J. T. E. (1972). "Nonparametric indexes of sensitivity and response bias," Psychol. Bull. 78, 429–432.
Richardson, J. T. E. (1976). "How to measure laterality," Neuropsychologia 14, 135–136.
Ryan, W. I., and McNeil, M. (1974). "Listener reliability for a dichotic task," J. Acoust. Soc. Am. 56, 1922–1923.
Shankweiler, D., and Studdert-Kennedy, M. (1967). "Identification of consonants and vowels presented to left and right ears," Q. J. Exp. Psychol. 19, 59–63.
Shankweiler, D., and Studdert-Kennedy, M. (1975). "A continuum of lateralization for speech perception?" Brain Lang. 2, 212–225.
Speaks, C., Niccum, N., and Carney, E. (1976). "Noninvariance of the ear advantage in dichotic listening," J. Acoust. Soc. Am. 60, S119(A).
Speaks, C., Niccum, N., Carney, E., and Marble, K. (1975). "Dichotic listening: Reversals in ear advantage," J. Acoust. Soc. Am. 58, S76(A).
Studdert-Kennedy, M., and Shankweiler, D. (1970). "Hemispheric specialization for speech perception," J. Acoust. Soc. Am. 48, 579–594.
Studdert-Kennedy, M., Shankweiler, D., and Pisoni, D. B. (1972). "Auditory and phonetic processes in speech perception: Evidence from a dichotic study," Cog. Psychol. 3, 455–466.
Tobias, J. V. (1972). "Curious binaural phenomena," *Foundations of Modern Auditory Theory*, edited by J. V. Tobias, Vol. II, pp. 463–486.
Van den Brink, G., Sintnicolaas, K., and Van Stam, W. S. (1976). "Dichotic pitch fusion," J. Acoust. Soc. Am. 59, 1471–1476.
Wexler, B. (1976). Personal communication.
Zangwill, O. L. (1960). *Cerebral Dominance in Relation to Psychological Function* (Oliver and Boyd, Edinburgh).