JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1991, 55, 125-132 NUMBER 1 (JANUARY)

QUANTITATIVE AND METHODOLOGICAL ASPECTS OF STIMULUS EQUIVALENCE

HENRY O'MARA

LISRONAGH, CLONMEL, IRELAND

The number of different ways of linking stimuli in the training phase of a conditional discrimination procedure designed to teach equivalence relations has hitherto been underestimated. An algorithm from graph theory that produces the correct number of such different ways is given. The establishment of equivalence relations requires transitive stimulus control. A misconception in a previous analysis of the conditions necessary for demonstrating transitive stimulus control is indicated. This misconception concerns responding in an unreinforced test trial to a negative rather than a positive comparison stimulus. Such behavior cannot be attributed to discriminative control by degree of association with reinforcement if the negative comparison stimulus has been less associated with reinforcement than the positive comparison stimulus in an antecedent training phase.

Key words: stimulus equivalence, transitive stimulus control, training cluster, graph theory, conditional discrimination

---

Assume that a subject takes part in a conditional discrimination procedure that has three phases. First, a probe phase shows that three stimuli (A, B, and C) bear no particular relation to one another so far as the subject is concerned. Second, in a reinforced training phase the following relation is taught: If A is present, then choose B rather than any other stimulus, where A is the sample stimulus (Sa), B is the positive comparison stimulus (Co+), and one or more other stimuli (not including C) are negative comparison stimuli (Co−). This relation is given the notation AB and another relation, BC, is taught in the same way. Third, an unreinforced test phase shows the existence of a new conditional relation, AC. This new relation is called a transitive relation.

A transitive relation cannot be called an equivalence relation unless, in addition to transitivity, the properties of reflexivity and symmetry are implicated (Sidman & Tailby, 1982). Reflexivity, symmetry, and transitivity are terms taken from mathematics, where they are the defining properties of an equivalence relation. Reflexivity is the matching of a stimulus to itself (e.g., AA, BB, CC). Symmetry is the reversal of a trained relation between a sample stimulus and a positive comparison stimulus (e.g., if AB, then BA). Transitivity, as indicated above, is the matching of two stimuli not directly related in reinforced training, but indirectly linked through their mutual association in trained relations with a third stimulus (e.g., if AB and BC, then AC). Strictly, transitivity further requires that the training relations that link the pair of stimuli in the emergent relation are set in the same direction. This means that, in the chain linking the pair, each connecting or mediatory stimulus must serve as a positive comparison stimulus in one trained relation and as a sample stimulus in another trained relation. Thus the example given above—if AB and BC, then AC—is a qualifying instance of transitivity, whereas the sequence—if AB and AC, then BC—is not. This does not mean that the emergent BC relation is not transitive, it just means that the trained relations AB and AC do not establish it as such. If AB is shown to be symmetric then BC is transitive, that is, the following qualifying sequence is established: if BA (established by symmetry with AB) and AC, then BC.

In addition to these formal conditions (the presence of a mediatory stimulus and the single direction requirement), the demonstration of transitive stimulus control demands a methodology that excludes extraneous or adventitious sources of control, or at least permits the

---

The author wishes to thank Leo Baker for unstinting support in the preparation of this paper. My gratitude is also due to Colm Ó Dúnlaing and Murray Sidman for their help. Correspondence and requests for reprints should be sent to the author at Lisronagh, Clonmel, County Tipperary, Ireland.

unconfounded assessment of their influence (see Fields, Verhave, & Fath, 1984, and the section of this paper titled, "Valence Disparity and the Assessment of Transitive Stimulus Control").

An alternative definition of stimulus equivalence is advocated by Vaughan (1988), who eschews the formal properties of reflexivity, symmetry, and transitivity in favor of a definition based on set partitions in mathematical set theory. This new definition is the subject of a controversy that is not considered here (see Hayes, 1989; Sidman, Wynne, Maguire, & Barnes, 1989; Vaughan, 1989).

ENUMERATION OF TRAINING CLUSTERS

In experimental settings, stimulus equivalence is usually investigated through the reinforced training of links (conditional relations) between members of a previously unrelated class of n stimuli and the subsequent unreinforced testing for emergent untrained links, especially transitive links, between the same stimuli. In the training phase, a minimum of $(n - 1)$ links is required; otherwise at least one of the stimuli will be omitted altogether.

The minimal training criterion of $(n - 1)$ links therefore incorporates the selection of an arrangement of training links that includes every stimulus in the class. But for all $n > 3$, this minimal training criterion can be satisfied by a plurality of arrangements of the interstimulus links. Each such unique arrangement is called a training cluster (TC) (Fields & Verhave, 1987). The identities of individual stimuli in training clusters are assumed to be arbitrary, in that switching stimuli while retaining the same arrangement of links does not create new training clusters. Thus, each of the two diagrams in Figure 1 represents the same training cluster. It is also assumed that training clusters are not differentiated by the direction of training of their constituent links, so that for a pair of stimuli, X and Y, that are linked in the training phase, the relative positions of X and Y (i.e., sample stimulus or positive comparison stimulus) are treated as being irrelevant. Given these assumptions (minimal training, arbitrariness of stimulus identities, and irrelevance of direction of training), it can be seen from Figure 2 that with a class of four stimuli a total of two training clusters is available.

As the number of stimuli increases, so too does the number of training clusters. Why is it important to know the number of training clusters for a class of a given size? The answer is that training cluster choice may be a factor in determining whether or how easily equivalence and transitivity are established. If so, experimenters will ignore potentially interesting choices insofar as their selections are based on incorrect, incomplete, or absent information on the number of training clusters available.

But what is the exact relation between n and TC? Fields and Verhave (1987) propose that it can be worked out as follows:

1. Identify n, the number of stimuli in the class.
2. For the particular n, identify the number of different ratios of nodes to singles, where, in the training phase, a node is a stimulus linked to more than one other stimulus, and a single is a stimulus linked to only one other stimulus. Thus, for a class of six stimuli, the following node/single ratios exist (a total of four): 1/5, 2/4, 3/3, and 4/2. In fact, for a class of n stimuli, there are always $(n - 2)$ node/single ratios.
3. For each node/single ratio, distribute the singles in all possible linking arrangements with the nodes. Figure 3 illustrates this for a class of six stimuli with a node/single ratio of 2/4.

Each distribution of singles to nodes at Step 3 above is a training cluster. These training clusters are summed first for each node/single ratio and then for the class as a whole. This gives the number of training clusters in the class, according to Fields and Verhave (1987), who provide the following formulas that produce the same results as their counting method just described:

for n odd,

$$TC = 2^{(n-4)} + 2^{(n-5)/2} \quad (1)$$

and for n even,

$$TC = 2^{(n-4)} + 2^{(n-4)/2}. \quad (2)$$

The counting method and its isomorphic formulas are flawed, however, and give understated TC values for all $n > 6$. This is because they do not count those training clusters in which a node or nodes are linked to

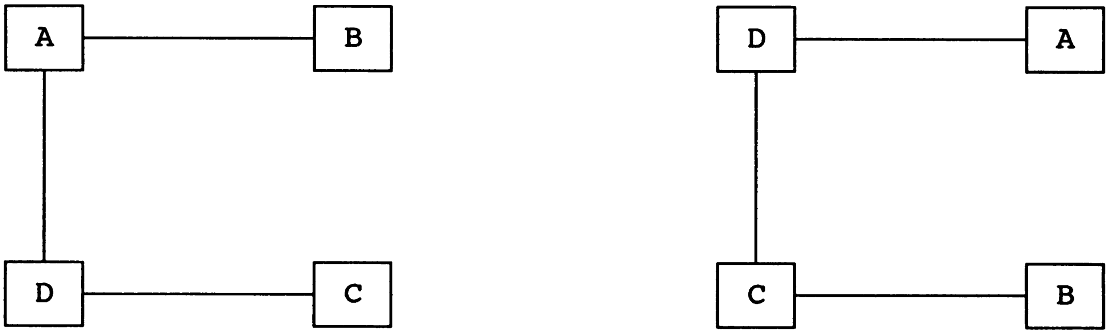

Fig. 1. Two arrangements of training links between the same stimuli constituting only one training cluster, given that the identities of individual stimuli are arbitrary for the purpose of differentiating training clusters.

more than two other nodes. The smallest class in which this occurs is $n = 7$ stimuli, for which the three-step counting method and Formula 1 predict 10 training clusters, whereas, in fact, there are 11 (see Figure 4). This undercounting is characteristic of the counting method and formulas of Fields and Verhave for all $n > 6$. As class size increases, instances of training clusters with nodes linked to more than two other nodes become more numerous.

To correctly obtain the number of training clusters from the size of the class requires a digression into that branch of mathematics known as graph theory. A graph may be loosely defined as a collection of points with lines joining some or all of them together. The points are called vertices and the connecting lines are called edges. A tree is defined as a connected acyclic graph, which means that (a) every vertex is joined to at least one other vertex and (b) there is only one path along the edges from any vertex to any other vertex. This second defining feature is conveyed more intuitively by saying that it is not possible to "go round in a circle" on the edges of a tree (see Figure 5). The number of edges in a tree is always one less than the number of vertices.

On reflection it becomes clear that the problem of finding the number of training clusters for a class of n stimuli is essentially the same as the problem of finding the number of trees with n vertices. Stimulus and vertex are corresponding terms, as are training link and edge. Recall the assumption made earlier that the identities of individual stimuli are arbitrary for the purpose of differentiating training clusters. In graph theory the parallel condition is that the vertices of a tree are indistinguishable, in which case the tree is described as free or unlabeled. The solution to the problem of enumerating free trees incorporates the enumeration of another type of tree described as rooted. In a rooted tree, one vertex (the root) is regarded as distinguishable from all the other vertices, which in turn are regarded as indistinguishable among themselves. To solve the problem of enumerating free trees let t_n and T_n denote, respectively, the number of free and

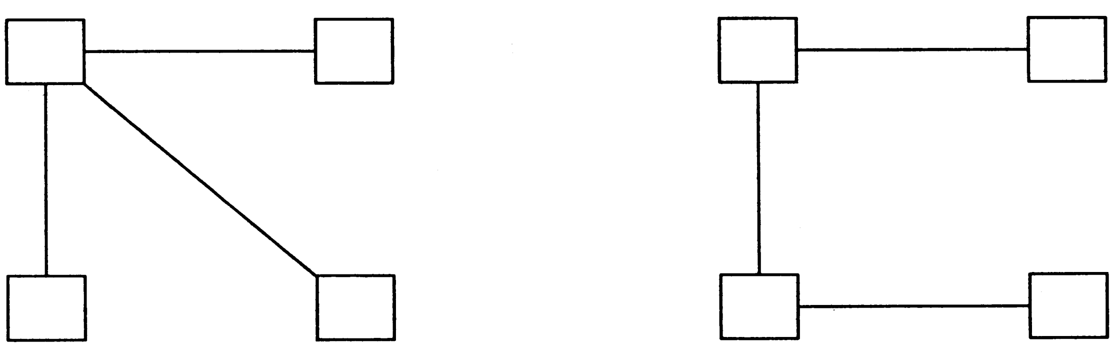

Fig. 2. All training clusters available for a class of four stimuli.

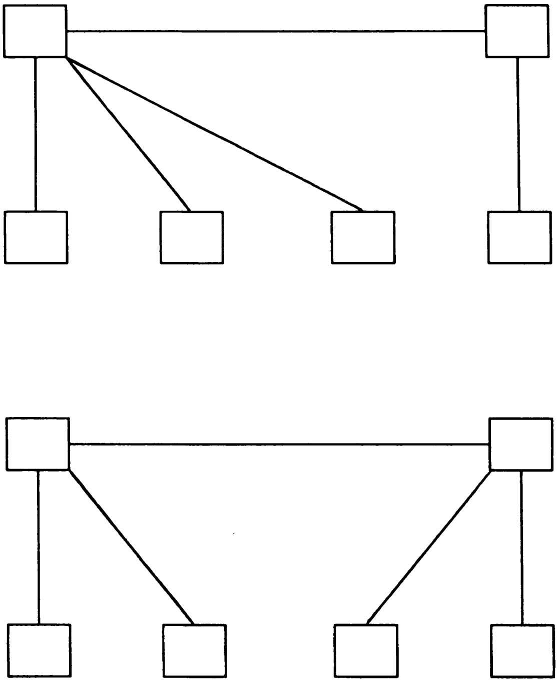

Fig. 3. Two of the training clusters for a class of six stimuli illustrating the different distributions available of four singles to two nodes.

rooted trees that have n vertices ($n = 1, 2, \ldots$). Using the method of generating functions, it follows that the series

$$t(x) = \sum_{n=0}^{\infty} t_n x^n \qquad (3)$$

and

$$T(x) = \sum_{n=0}^{\infty} T_n x^n, \qquad (4)$$

where these are infinite series whose terms involve powers of a dummy variable x.

Given these premises it can be shown that

$$t(x) = T(x) - \tfrac{1}{2}x[T^2(x) - T(x^2)]. \qquad (5)$$

This generating function or counting series was first discovered by Otter (1948). An alternative and shorter proof is given by Clarke (1959), on whose exposition this presentation is heavily reliant. Another useful treatment is by Deo (1974), whose elaboration of Formula 5 for the first 10 terms gives

$$t(x) = x + x^2 + x^3 + 2x^4 + 3x^5 + 6x^6 + 11x^7 + 23x^8 + 47x^9 + 106x^{10} + \ldots . \qquad (6)$$

The coefficient of x for the nth term of this series gives the number of free trees with n vertices and also the TC value for a class of n stimuli. The series can be extended further; the results for n values between 3 and 22 are given in Table 1.

The following asymptotic formula, adapted from Otter (1948), can be used to estimate the number of free trees/training clusters (t_n) for large n:

$$t_n = \frac{0.5349485}{n^{5/2}\alpha^n} + \epsilon_n, \qquad (7)$$

where $\alpha = 0.3383219$ and ϵ_n (the error term) is small compared to the first term when n is large. It is clear from Table 1 and from the application of the asymptotic formula that the number of training clusters becomes very large as class size increases.

It was assumed earlier that the identities of individual stimuli are arbitrary for the purpose of differentiating training clusters. If this assumption is removed completely and each stimulus is treated as unique (implying that the two diagrams in Figure 1 actually represent different training clusters), then the number of training clusters is given by the following equation:

$$\text{TC} = n^{(n-2)}. \qquad (8)$$

Formula 8 is again taken from graph theory, where the comparable problem is the enumeration of labeled trees. It was first stated by Cayley (1889), and an accessible proof is given by Deo (1974). As would be expected, the TC values produced by this relation are much greater than those obtained when individual stimulus identities are deemed arbitrary. The results of Formula 8 for n values from 3 to 10 are given in Table 2.

A further assumption was that the direction of training of a link between a pair of stimuli is not a factor in differentiating training clusters. If this restriction is done away with, then the number of training clusters must increase.

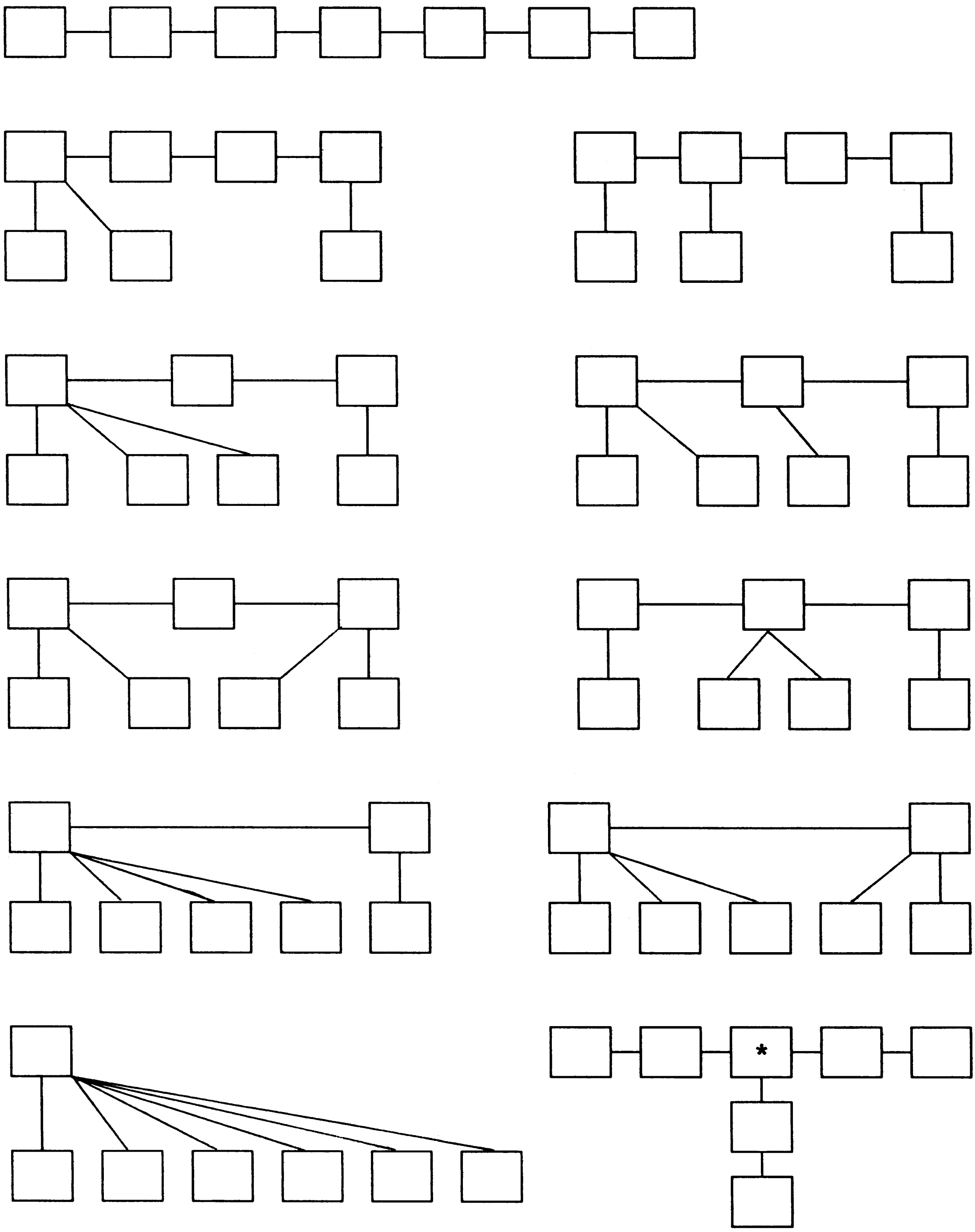

Fig. 4. All training clusters available for a class of seven stimuli. The training cluster at the bottom right contains one node (marked "*") that is linked to three nodes and to no singles. This is the "extra" training cluster not counted by the analysis of Fields and Verhave (1987).

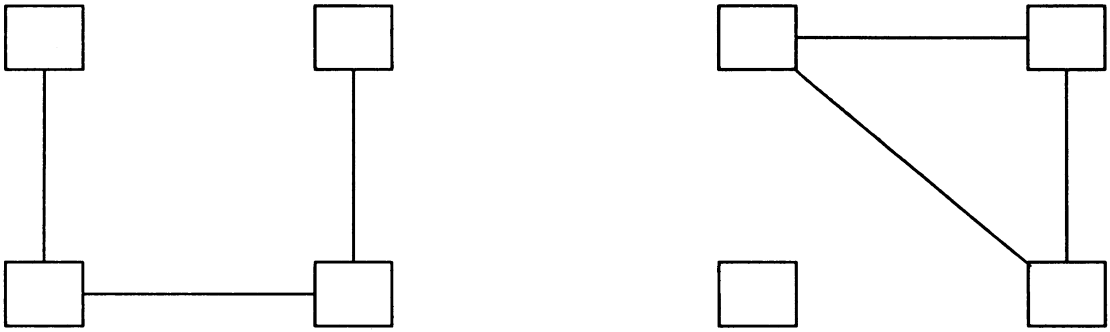

Fig. 5. The graph on the left is a tree, formally a connected acyclic graph, whereas the graph on the right is not a tree because it contains a cycle.

In these circumstances, the exact relation between n and TC constitutes a problem that is dealt with by graph theory as a "coloring" problem. This is a technical and involved area and is not pursued here.

VALENCE DISPARITY AND THE ASSESSMENT OF TRANSITIVE STIMULUS CONTROL

The hypothetical subject cited earlier was taught the relations AB and BC and then, in an unreinforced test phase of the procedure, demonstrated the AC relation. AC was assumed to be transitive, that is, it was a function of the pairing of both Stimuli A and C with the third Stimulus B. But in such settings AC is not necessarily transitive. Its appearance may instead reflect the action of some other source of control. Conditions under which such nontransitive control may operate have been investigated by Fields et al. (1984). They propose as a paradigmatic case the establishment of two equivalence classes with minimal training, and with the stimuli of each class serving as negative comparison stimuli for training and test trials in which stimuli from the other class are presented as sample and positive comparison stimuli. All trials contain only three stimuli: Sa, Co+, and Co−.

With this design, one possibility for the confounding or subversion of transitive stimulus control involves the "valence" of the comparison stimuli. As defined by Fields et al. (1984), the valence of a comparison stimulus is the frequency with which it is used as a Co+ during training minus the frequency with which it is used as a Co− during training. The relative valence of the comparison stimuli in a test trial may render the assessment of transitive stimulus control problematic or impossible, because responding during test trials may

Table 1

Numbers of training clusters (TC) for stimulus classes (n) containing 3 to 22 stimuli.

n	TC
3	1
4	2
5	3
6	6
7	11
8	23
9	47
10	106
11	235
12	551
13	1,301
14	3,159
15	7,741
16	19,320
17	48,629
18	123,867
19	317,955
20	823,065
21	2,144,505
22	5,623,756

Table 2

Numbers of training clusters (TC) for stimulus classes (n) containing 3 to 10 stimuli in which the identities of individual stimuli are not arbitrary for the purpose of differentiating training clusters.

n	TC
3	3
4	16
5	125
6	1,296
7	16,807
8	262,144
9	4,782,969
10	100,000,000

come under discriminative stimulus control rather than transitive stimulus control, inasmuch as behavior may be governed by the rule, "select whichever comparison stimulus has the higher valence." Three types of test trials are distinguished on the basis of the valence disparity of the comparison stimuli. In a "strong" trial the valence of the Co+ is less than the valence of the Co−; in a "neutral" trial their valences are equal; and in an "inadequate" trial the valence of the Co+ is greater than that of the Co−.

In relation to inadequate test trials, Fields et al. (1984) state ". . . an inadequate test configuration would not be helpful in distinguishing between discriminative stimulus control [i.e., control by valence disparity] and transitive stimulus control if responding occurred to Co+. Responding to Co−, however, would provide strong evidence for discriminative control and against transitive stimulus control" (p. 153). This statement is correct insofar as Co+ responding is concerned. It is also correct in saying that Co− responding with an inadequate test configuration provides strong evidence against transitive stimulus control. What is not correct, however, is the parallel assertion that Co− responding with an inadequate test configuration provides strong evidence for discriminative control. If the valence disparity of the comparison stimuli in that configuration is in fact exercising discriminative control, then behavior will agree with the rule, "select whichever comparison stimulus has the higher valence." But where inadequate test trials are used, such behavior must constitute the choice of Co+, not Co−. It is clear, therefore, that responding to Co− in an inadequate test trial setting is evidence against discriminative control by valence disparity and is also evidence against control by transitivity. Such responding requires explanation by reference to some other locus of control. Of course, inadequate test trials, as their name indicates, should not be used to investigate the establishment of transitive stimulus control. They are defective in that, where Co+ responding occurs, they confound the effects of transitive stimulus control and of discriminative control by valence disparity.

DISCUSSION

Though largely emendatory, this paper may have some function in highlighting two aspects of stimulus equivalence of interest to experimenters.

First, it points to a previously underestimated repleteness of options, in terms of particular combinations of relations between stimuli, that is available for presentation in training procedures. This raises the possibility that, where equivalence or transitivity cannot be demonstrated with a particular training cluster, recourse to another training cluster might be successful. Of course, this possibility may not exist. The training cluster chosen may have no special influence on the actuality or the ease of establishment of equivalence or transitivity; the experimental literature to date has not addressed this question. Future empirical studies could compare all training clusters for classes with small numbers of stimuli. With larger class sizes this would be impractical, and only a proportion of the available training clusters could be used. One selection criterion could be based on "associative distance," which is defined by Fields et al. (1984) as the number of training relations joining two stimuli in a putatively transitive pair. Associative distance measurements could be summed for all such pairs to give an overall index for the training cluster. Then training clusters with different associative distance index values could be compared.

Second, the preceding section of the paper emphasizes the importance of a rigorous and predetermined selection of those stimulus combinations to be presented in the training and test phases of experiments. Where the design neglects this desideratum, it may be impossible to infer from the results whether equivalence and transitivity are or are not manifest.

REFERENCES

Cayley, A. (1889). A theorem on trees. *Quarterly Journal of Mathematics*, **23**, 376–378.

Clarke, L. E. (1959). On Otter's formula for enumerating trees. *Quarterly Journal of Mathematics*, Oxford Ser. 2, **10**, 43–45.

Deo, N. (1974). *Graph theory with applications to engineering and computer science*. Englewood Cliffs, NJ: Prentice-Hall.

Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of the Experimental Analysis of Behavior*, **48**, 317–332.

Fields, L., Verhave, T., & Fath, S. (1984). Stimulus equivalence and transitive associations: A methodological analysis. *Journal of the Experimental Analysis of Behavior*, **42**, 143–157.

Hayes, S. C. (1989). Nonhumans have not yet shown stimulus equivalence. *Journal of the Experimental Analysis of Behavior,* **51,** 385–392.

Otter, R. (1948). The number of trees. *Annals of Mathematics,* 2nd Ser., **49,** 583–599.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior,* **37,** 5–22.

Sidman, M., Wynne, C. K., Maguire, R. W., & Barnes, T. (1989). Functional classes and equivalence relations. *Journal of the Experimental Analysis of Behavior,* **52,** 261–274.

Vaughan, W., Jr. (1988). Formation of equivalence sets in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes,* **14,** 36–42.

Vaughan, W., Jr. (1989). Reply to Hayes. *Journal of the Experimental Analysis of Behavior,* **51,** 397.

Received October 3, 1989
Final acceptance September 11, 1990