



Toward a theory of variability discrimination: finding differences

M.E. Young^{a,*}, M.R. Ellefson^a, E.A. Wasserman^b

^a Department of Psychology, Southern Illinois University, Carbondale, IL 62901-6502, USA

^b Department of Psychology, University of Iowa, Iowa City, IA 52242, USA

Received 26 June 2002; accepted 11 October 2002

Abstract

We sketch the outlines of a theory of variability discrimination that aggregates localized differences to mediate variability discrimination. This Finding Differences Model was compared to a Positional Entropy Model across four different data sets. Although the two models provide strong and similar fits across three of the data sets, only the Finding Differences Model is applicable to investigations involving multidimensional variability. Furthermore, the Finding Differences Model is based on an activation map that has been shown to have utility for visual search tasks, thus establishing its generality across task domains. © 2003 Elsevier Science B.V. All rights reserved.

Keywords: Variability discrimination; Same–different; Entropy; Modeling

1. Introduction

People judge variety in a range of everyday tasks. We judge the variety of products on a store shelf, racial diversity in our work environment, and restaurant fare in our community. Psychologists have documented that people have preferences for certain levels of variety (Berlyne, 1960; Munsinger, 1966; Munsinger and Kessen, 1966) and marketing researchers have reported that perceived variety has a strong effect on decisions of where to shop (e.g. Baumol and Ide, 1956; Broniarczyk et al., 1998; Hoch et al., 1999). Despite the everyday relevance of variability discrimination, only recently have researchers begun to examine whether non-human animals can judge variety and how they might do so.

Young, Wasserman, and co-workers have extensively studied variability discrimination in pigeons,

baboons, and people (e.g. Wasserman et al., 1995, 2000, 2001; Young and Wasserman, 1997, 2001, 2002; Young et al., 1997b). All three species respond discriminatively to stimuli involving different degrees of variability. Fig. 1 shows two examples of typical displays used in that research. These displays exhibit the minimal and maximal degrees of variability possible in collections of 16 items. Not only do all three species respond differently to these two endpoints of the variability dimension, but they respond systematically to intermediate values as well.

Young and Wasserman (1997) found that the information theoretic measure of categorical variability, *entropy*, nicely captures the functional relationship between stimulus variability and discriminative responding. To quantify entropy, Shannon and Weaver (1949) used the following equation:

$$H(A) = - \sum_{a \in A} p_a \log_2 p_a \quad (1)$$

where $H(A)$ is the entropy of categorical variable A , a is a category of A , and p_a is the proportion of observed

* Corresponding author. Tel.: +1-618-453-3567.
E-mail address: meyoung@siu.edu (M.E. Young)

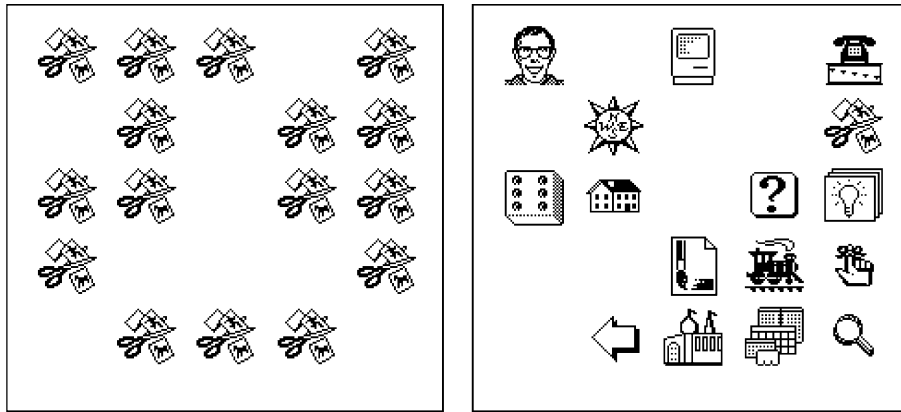
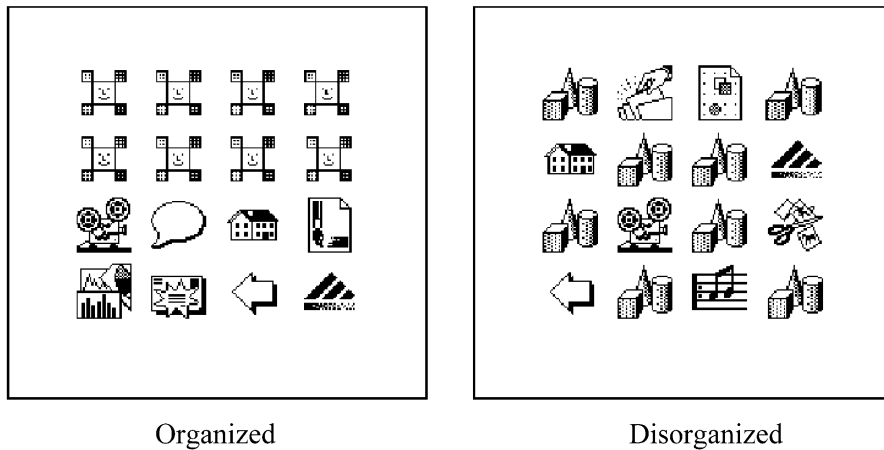


Fig. 1. Examples of the 16-icon arrays used during training in many of the studies by Young, Wasserman, and co-workers. These arrays consisted of 16 icons chosen from a set of 24 to create Same and Different arrays with these 16 icons randomly located in a 5×5 grid.



Organized

Disorganized

Fig. 2. Examples of the Organized displays (identical items clustered together) and Disorganized displays (identical items widely distributed) used in Wasserman et al. (2000). These arrays involved eight identical items and eight different items.

values within that category. When a display has 16 identical icons, there is only one category (the single icon) with a probability of occurrence of 1.0. Because $\log_2(1.0) = 0.0$, the entropy of a same display is 0.0. A different display consists of one occurrence of each of 16 icons (i.e. 16 shape categories), yielding an entropy of $-0.0625 \times \log_2(0.0625) \times 16$, or 4.0. Displays with intermediate degrees of variability have intermediate levels of entropy.

Although entropy has provided an excellent initial fit of the pigeons' discriminative performance, the modulating effect of item location was documented in a later study involving a positional variant of entropy

(Wasserman et al., 2000). When identical items were clustered together (left side of Fig. 2), pigeons were slightly more likely to choose the "same" key than when identical items were more widely distributed (right side of Fig. 2).¹ Wasserman et al. (2000) provided a measure of positional entropy in which individual entropy scores were calculated for all possible contiguous subsets (e.g. all nine possible groups of contiguous items in a 2×2 arrangement) and subsequently averaged. This positional variant of entropy

¹ We conducted an analogous study involving humans and found very similar results Young and Ellefson (2002).

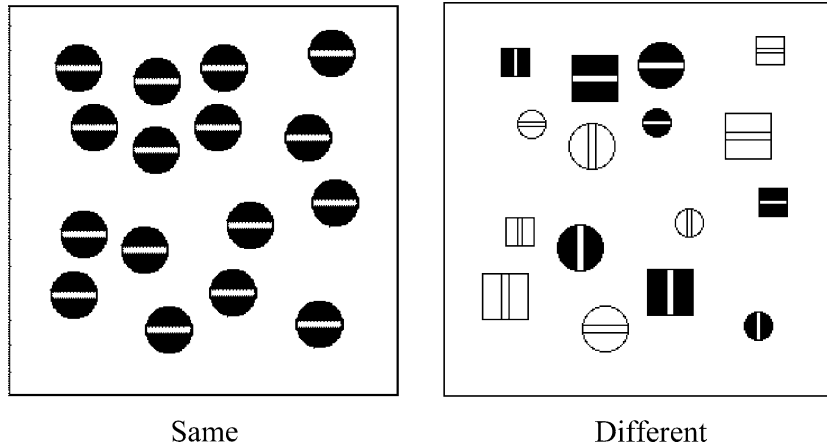


Fig. 3. Examples of the Same and Different displays used in Young et al. (2002a). These arrays involved up to 16 possible items created from all possible combinations of two levels of brightness, size, orientation, and shape ($2 \times 2 \times 2 \times 2 = 16$).

accounted for the effect of the organization of same and different items in mixtures like those shown in Fig. 2.

In even more recent work, Young and co-workers have documented that the degree of difference is important to performance, not just whether two items are identical or non-identical. In nearly every published study to date, the multiple-stimulus displays have involved variability in shape—brightness and approximate size have been constant. Two new studies involving pigeons (Young et al., 2002a) and humans (Young and Ellefson, 2002) used displays like those shown in Fig. 3, in which the items varied in shape, brightness, size, and orientation. After receiving training requiring discriminative responding to same and different arrays, participants were tested with four different types of displays in which each of the four dimensions was held constant (e.g. a brightness-constant array involved 16 white stimuli which varied in size, shape, and orientation or 16 black stimuli which varied in size, shape, and orientation). The salience of the dimension importantly determined the response profile of the subjects. When brightness was held constant (apparently a very salient dimension in our task), the proportion of “same” responses increased dramatically, but when orientation was held constant (apparently a much less salient dimension), responding was largely unaffected. Young and Ellefson (in press) also reported independent control of variability in shape and color

using chromatic versions of the icons shown in Fig. 1.

The effects of judged similarity produce further problems for the original entropy account. Entropy, as calculated by Eq. (1), requires that items are either in the same category or not—there is no room for degree of category membership. To account for these new data from an entropy perspective, a measure of multidimensional entropy would be required.

2. Toward a new theoretical account: finding differences

Although we could continue to search for extensions of entropy (e.g. posit a multidimensional, positional entropy), we opted to approach the problem from a different perspective. Perhaps variability discrimination leverages stimulus information that has proven to be useful in performing other tasks. Rather than proposing an altogether new mechanism that is specifically dedicated to detecting variability, we considered the possibility that a re-description of the stimulus that has been used in the service of a much more common task, visual search, might provide the basis for variability discrimination.

In a typical visual search task, a rich array of items is presented: one of which is the target of the search and the rest of which are distractors. For example, the participant might receive instructions to find a T

among Ls or a red X among red Os and blue Xs (e.g. Treisman and Gelade, 1980). The similarity of the target to the distractors has an important effect on the efficiency of the search: the more similar the target is to the distractors, the harder it is to find (Duncan and Humphreys, 1989).

Wolfe (1994) proposed that each item in the array is “active” to various degrees; an item’s level of activation establishes the degree to which it stands out from its neighbors thus determining the difficulty of finding it during a search task. In his model, Guided Search 2.0 (GS2), each item receives two sources of activation: bottom-up and top-down. Bottom-up activation is produced by the array items themselves; those items that are different from their neighbors will be more active (i.e. stand out). Top-down activation is produced by expectations usually induced by task demands or instructions. For example, if an observer is reinforced for finding red items or is told to search for a red item, then color will receive extra attention; greater attention to color (a top-down influence) will make bottom-up color differences more salient.

In GS2 (Wolfe, 1994), the bottom-up sources of activation are captured in an activation map. This map includes an activation score for each item in the display. Those items that differ from their neighbors receive greater activation than those that are similar to their neighbors. This re-description of the original display thus captures *local differences*. For a visual search task, the location of salient differences helps to direct attention to those items that stand out from the background.

In our examination of this theory, we noted that the average activation in GS2’s activation map provides a plausible index of the degree of variability in the array. If the map contains many highly active items, then the display must comprise many items that are different from their neighbors (i.e. it involves high variability). If the map contains few active items, then the display must comprise items that are very similar to one another (i.e. it involves low variability). We were attracted to the parsimony of using Wolfe’s activation map for both visual search (in which the locations of differences are important) and variability discrimination (in which only the average level of activation is important). No additional “entropy detection” mechanism need be posited.

This approach has three other potential benefits as a theoretical account of performance in our tasks. First, Wolfe’s activation map computes activation based on local differences thus providing a possible account of the effect of organization on variability discrimination (see Fig. 2). Second, the similarity of an item to its neighbors is computed in producing the map, thus providing a mechanism for the degree of similarity to affect discrimination performance (see Fig. 3). Third, the role of top-down processes in GS2 allows the modeling of differential dimensional salience produced by perceptual or task demands (e.g. if participants were reinforced for responding to color variability but not to shape variability, then differences in color should be emphasized, thereby producing the greatest activation for those items that differ in color from their neighbors, Young and Ellefson, in press).

Given the promise of GS2’s activation map as a basis for variability discrimination, we are developing a new model that can capture these functional relationships (Young et al., 2002b). Here, we explore whether the model can account for variability discrimination performance in the pigeon and compare this *Finding Differences Model* to the *Positional Entropy Model* reported by Wasserman et al. (2000).

2.1. Finding Differences Model

The activation maps used in our simulations were either 4×4 or 5×5 matrices, each value of which represented the activation (i.e. “differentness”) of a particular item in the original display. We are not proposing that such an array actually exists in the head of the observer; the activation map merely represents a computational re-description of the display that provides significant utility in our model of the functional relationships between the properties of the stimulus and the behavior of the observer.

The activation of an item is a function of two factors: (a) the degree of difference between it and each of its neighbors and (b) the distance between it and each of its neighbors. Rather than define an item’s neighborhood as only those items that are adjacent to or within some pre-specified distance, we scaled the degree of difference between items as a function of their Euclidean distance on the screen. The computational details and other aspects of the complete model can be found in Young et al. (2002b).

To measure the degree of difference between items, we relied on Shepard’s (1964) Minkowski metric. We adopted the city-block version, because our items were composed of separable dimensions (Shepard, 1991). For our tasks, the difference between stimuli was thus the sum of their feature differences. In the simulations, we assumed that differences in size, brightness, shape, or orientation each produced a difference score of 1.0. For example, when comparing a large, black square with a vertical bar to a small, black circle with a vertical bar (see Fig. 3 for examples), the resulting difference score, d , is 2.0. These two stimuli thus would be judged to be more similar than the same large, black square with a vertical bar is to a small, white circle with a horizontal bar (difference score, d , of 4.0). The model allows for degrees of difference within each dimension (a white circle would be more similar to a gray circle than to a black circle), but we have yet to examine this issue empirically or theoretically.

To incorporate differences in salience among the dimensions (due to bottom-up or top-down influences), each dimension’s influence on similarity was scaled as a function of its salience by simply differentially weighting any computed feature differences in the Minkowski metric (e.g. differences in orientation would have less of an impact on the computation of stimuli differences than would differences in brightness). Differences in salience may be due to prior

reinforcement contingencies, the perceptual abilities of the observer, or the instructions provided to the observer (in human studies).

These difference scores are aggregated for an item and scaled by the Euclidean distance between the item and its neighbors to produce an activation score for each item in the display. We used Euclidean distance to reduce the effect of any differences as the distance between items increased. Although we modeled this relationship as exponential, we have not systematically explored whether other functions might better capture the true relationship between distance and judged differentness.

After the original display was re-described as an activation map (see Fig. 4 for an example stimulus array and its corresponding activation map), we simply averaged across the activation values in the map to produce an aggregate measure of the differences that were present in the array, \bar{a} . This measure could then be mapped to a response through the reinforcement contingencies of the task. Rather than modeling this learning process, we represented the end-state of learning a two-alternative forced choice task as a sigmoidal relationship:

$$P(R = \text{different}) = (1 + e^{-m(\bar{a}+b)})^{-1} \tag{2}$$

in which m and b are free parameters that determine the sharpness of the gradient and the

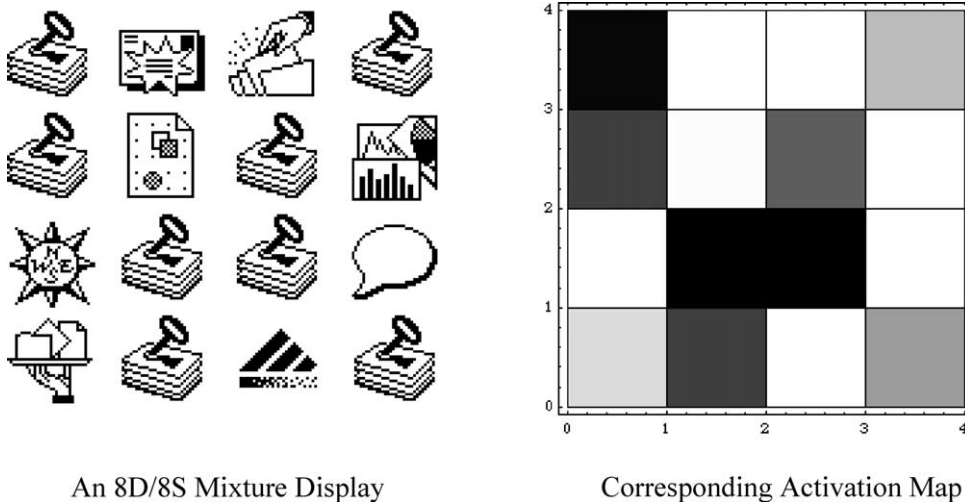


Fig. 4. An example array (8D/8S designates a display with 8 different items and 8 same items) and its corresponding activation map in which white designates high activation, black designates low activation, and shades of gray designate intermediate levels of activation.

location of the psychological midpoint of the scale, respectively.

2.2. Fitting the data

In comparing two models, it is important to consider two factors: the fit of the model to the data and the parsimony of the model (Myung, 2000; Pitt et al., 2002). The basic entropy formula contains no free parameters, a remarkably parsimonious model. However, in order to fit data showing an effect of spatial organization (Wasserman et al., 2000), a “window of attention” was introduced with two free parameters, the *height* and *width* of the window, to produce the Positional Entropy Model. The Finding Differences Model captures the effect of localized differences through the use of a single parameter, c , that scales the influence of distance on similarity (Young et al., 2002a,b). The model also includes a set of additional free parameters for the salience of each dimension. Salience is only relevant for the experiments involving multiple dimensions of variability and therefore is not used (i.e. it is held constant) for most of our simulations. A new, multidimensional entropy formula would be necessary, however, for experiments involving multiple dimensions of variability and hence would require its own set of unspecified parameters.

Finally, Eq. (2) represents the mapping of the extracted variability to an actual response. It involves two parameters: m and b . In our prior fits of entropy to data, two parameters have also been necessary, the *slope* and *intercept* of a regression equation.

In sum, any observed advantage of the Finding Differences Model over the Entropy Model cannot be due to its possessing more degrees of freedom. Indeed, the Positional Entropy variation of Wasserman et al. (2000) has a free parameter not found in the new model. In this initial examination of the relative utility of the Positional Entropy and Finding Differences Models, we will focus our attention on fitting the models to a series of data sets obtained in studies of variability discrimination by pigeons. Although researchers have found species differences (e.g. Wasserman et al., 2001; Young and Wasserman, 2002), an exploration of these differences will not be considered here.

Four results will be examined: (a) as display variability increases, pigeons are more likely to choose the

“different” key (Young and Wasserman, 1997, 2002), (b) when identical items are clustered together, pigeons report the display to be less variable than when the items are spread out (Wasserman et al., 2000), (c) discriminative performance decreases as the number of items decreases, but only for Different displays (Young et al., 1997b), and (d) when the display comprises multidimensional items of the sort shown in Fig. 3, as the items are made more similar (e.g. by holding brightness constant), pigeons are more likely to judge the display as “same” (Young et al., 2002a).

Because the spatial organization effects of Wasserman et al. (2000) represent the only published data requiring the Positional Entropy Model to account for the results, we first ran simulations of that study to determine the optimal c value (representing the effect of spatial distance). Wasserman et al. reported data showing that a 2×3 attentional window provided the best fit for the Positional Entropy Model. Subsequent simulations used these optimal values to fit data from the other studies; we did not assume that the effect of distance varied across studies.

2.2.1. Spatial organization

The first set of simulations involved a systematic exploration of the optimal value of c that is necessary to fit the data from Experiment 2 of Wasserman et al. (2000). To remove individual biases toward choosing one or the other report key, the pigeon data were standardized so that the Same arrays were given a score of 0 and the Different arrays a score of 1. These arrays were not used in finding the optimal c , m , and b values, but they were included in assessing the fit of each model.

We ran a series of simulations across a range of c values from 0.2 to 1.0 in steps of 0.1, optimizing the m and b parameters of Eq. (2) using gradient descent for each run.² Smaller c values indicate little effect of spatial distance, whereas larger c values indicate

² Our software would not converge when we tried to use gradient descent to simultaneously find the optimal values for c , m , and b . Given that the fits were least sensitive to changes in c , we opted for the approach described here. For the other data sets, smaller c values sometimes produced better fits, whereas larger c values (especially greater than 1.0) invariably produced poor fits. It is possible that the effect of distance changes when the items are placed in 5×5 arrays rather than in the 4×4 arrays used by Wasserman et al. (2000).

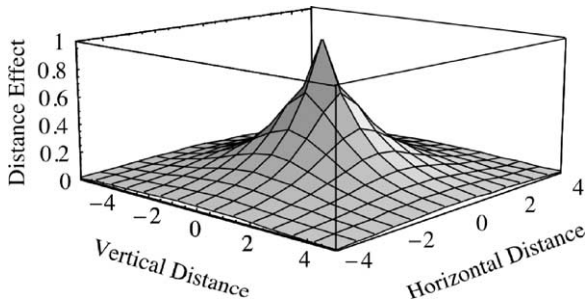


Fig. 5. The effect of spatial distance on similarity when c is 0.65. For an icon located in the center of this figure, the effect of a difference between it and a distant icon decreases as the distance between the icons increases. Distance was measured based on row and column position. Thus, adjacent items in the same row or column are 1.0 unit apart, and diagonally adjacent items are 1.4 units apart. The greatest possible distance in our 5×5 arrays was between items at opposite corners, 5.7 units. We used positive and negative distance values in the figure to designate left/right and above/below.

greater effects of spatial distance. The optimal value for c was 0.65; Fig. 5 illustrates the effect of spatial distance for this value.

Both the Positional Entropy and Finding Differences Models provided excellent fits of the data, $R^2 = 0.99$ and 0.98 , respectively. Each captured the overall effect of clustering the identical items together (Organized displays); these displays were rated as more “same” than those displays in which the same collection of items involved no clustering (Disorga-

nized displays, see Fig. 6). Although small details of the pigeon data were not fully captured by either model, the overall fits were very good.

2.2.2. Sensitivity to mixtures

The original study finding an effect of item mixture involved a number of different types of icon mixtures across three different experiments (Young and Wasserman, 1997). A more recent study involved 11 different mixtures that spanned an even larger entropy range than the original study within a single experiment (Young and Wasserman, 2002). We used the data from the more recent study to test the generality of the models. As before, the pigeon data were standardized so that the Same arrays were given a score of 0 and the Different arrays a score of 1. These arrays were not used in finding the optimal m , and b values, but they were included in assessing the fit of each model (i.e. as an extrapolation test).

The displays were 5×5 arrays in which 9 of the 25 locations were empty (and thus were not involved in the calculations except by altering the distance between items). Additionally, because the location of the items within these arrays was randomized in the Young and Wasserman (2002) study, we used 10 different randomizations of each type of mixture (a total of 110 arrays) to choose the optimal m and b parameter values. A different 10 randomizations of each type of mixture were later used to produce predictions from each model using these optimal

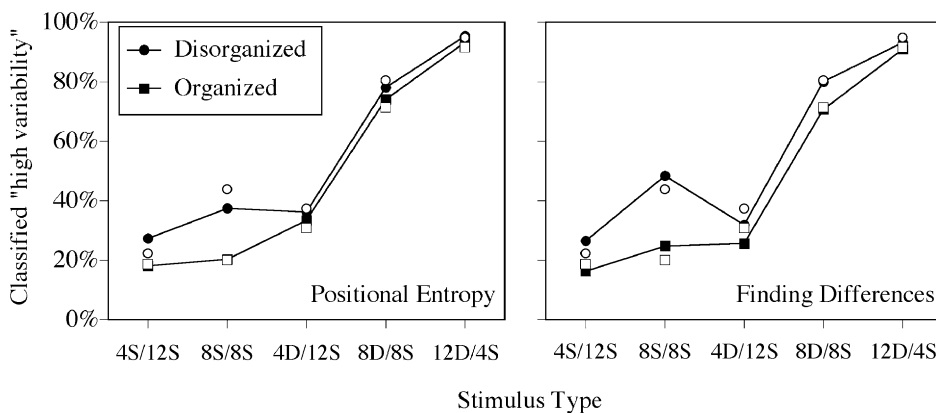


Fig. 6. The simulation results for the Organization study. The mS/nS and mD/nS notation indicates the type of mixture (see Wasserman et al., 2000, for details). The plots show the best fitting Positional Entropy and Finding Differences Models (solid symbols connected by lines) and the actual outcome from Wasserman et al (open symbols). The optimal m and b parameters for the models were -1.18 and 2.47 , and 0.123 and 23.51 , respectively.

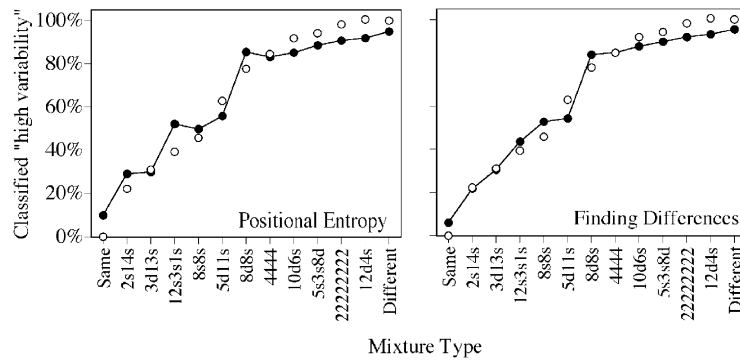


Fig. 7. The simulation results for the Mixtures study. The X-axis labels indicate the type of mixture sorted by their variability as judged by the pigeons (see Young and Wasserman, 2002, for details). The plots show the best fitting Positional Entropy and Finding Differences Models (solid symbols connected by lines) and the actual outcome from Young and Wasserman (2002) (open symbols). The optimal m and b parameters for the models were -0.79 and 2.78 , and 0.071 and 62.46 , respectively.

parameter values. These predicted scores were averaged to produce a single predicted score for each mixture.

Both Positional Entropy and Finding Differences Models again provided excellent fits of the data, $R^2 = 0.97$ and 0.98 , respectively, notably higher than the fit provided by the non-positional entropy of the arrays, $R^2 = 0.92$; each of the two focus models captured the overall effect of increasing stimulus variability (Fig. 7). These fits were very strong given that the value of c and the window size was found using a different set of data.

2.2.3. Number

The next set of simulations involved fitting the data from Experiment 1 of Young et al. (1997b) which investigated the effects of varying the number of items on Same and Different trials. We did not standardize the scores using the 16-icon Same and Different arrays because no key bias was apparent. Given the small numbers of data points for each display type (2, 4, 8, 12, and 16 icons for Same and Different displays), we opted for the additional point along each function. The simulation procedure was otherwise identical to that used for the mixture simulations, because Young et al. (1997b) also used 5×5 arrays in which the icons were randomly placed.

Both Positional Entropy and Finding Differences Models provided excellent fits of the data, $R^2 = 0.97$ and 0.94 , respectively, again higher than the fit provided by the non-positional entropy of the arrays,

$R^2 = 0.91$; each of the two focus models captured the overall asymmetry of the effect of item number on Same and Different arrays (Fig. 8), although each of the models had problems making accurate predictions for the 8-icon Different displays. These fits again were very strong given that the value of c and the window size was found using a different set of data.

2.2.4. Relative similarity (four-dimensional studies)

In a recent pair of experiments, Young et al. (2002a) trained pigeons to discriminate the variability of displays like those shown in Fig. 3. When each of the dimensions was held constant during a generalization test, the pigeons responded as a function of the constant dimension. When brightness was constant, subjects were much more likely to choose the “same” key than when it varied; when size was constant, they were slightly more likely to choose the “same” key than when it varied; and, when shape or orientation was held constant, their behavior was largely unaffected. Thus, their behavior was primarily controlled by differences in brightness with little effect of the other dimensions. These results are problematic for the original entropy account, because the entropy for all of the displays involving a constant dimension is 3.0. The Positional Entropy Model falls prey to the same problem, because entropy is based on identity versus non-identity—there is no room for degrees of similarity.

We applied only the Finding Differences Model to the data of Young et al. (2002a), because we have been

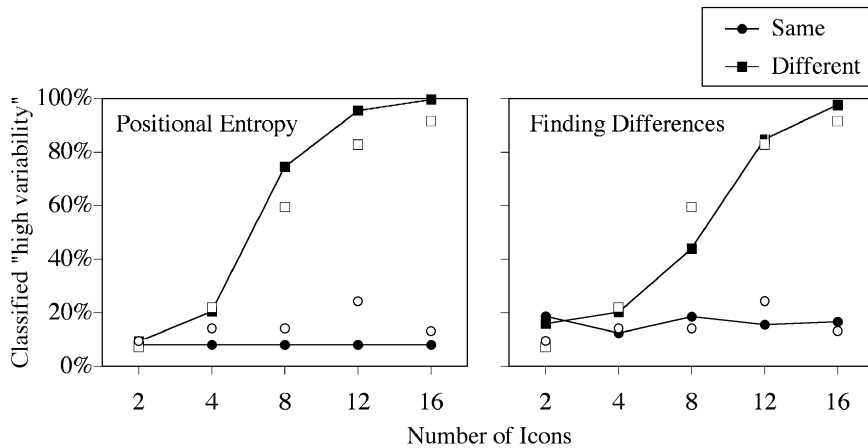


Fig. 8. The simulation results for the Number study. The plots show the best fitting Positional Entropy and Finding Differences Models (solid symbols connected by lines) and the actual outcome from Young et al. (1997b) (open symbols). The optimal m and b parameters for the models were -0.49 and 5.00 , and 0.070 and 80.82 , respectively.

unsuccessful in identifying a multidimensional analogue of entropy that allows for differential salience of the requisite dimensions. These simulations used salience parameters of the model to incorporate dimension differences. Each item in an array was coded as a vector of 0s and 1s, with 0 assigned to one of the values of the dimension (e.g. large) and 1 assigned to the other (e.g. small). For example, if a large, black square with a vertical bar was represented as $\{0, 0, 0, 0\}$, then a small, white circle with a vertical bar was represented as $\{1, 1, 1, 0\}$. Given the small numbers of data points and the possible problem with overfitting, we did not attempt to find the optimal salience values, but instead estimated them by using the empirically derived Cohen’s d effect sizes from Experiment 1 of Young et al. (2002a): 1.53, 0.10, 0.06, and 0.01, for brightness, size, orientation, and shape, respectively. The procedure was otherwise identical to that used in the mixture and number simulations.

The Finding Differences Model provided an excellent fit of the data, $R^2 = 0.997$; but, interpreting this high correlation must be qualified because only very low and very high response probabilities were observed, possibly inflating the calculated correlation (Fig. 9). Be that as it may, the simulations clearly indicated that the model can capture the qualitative relationship between similarity—as mediated by differential dimensional salience—and discriminative performance on a variability task.

2.2.5. Summary

For three of the four data sets, both focal models performed very similarly, accounting for most of the variance in mean discriminative performance. The effect of spatial distance established

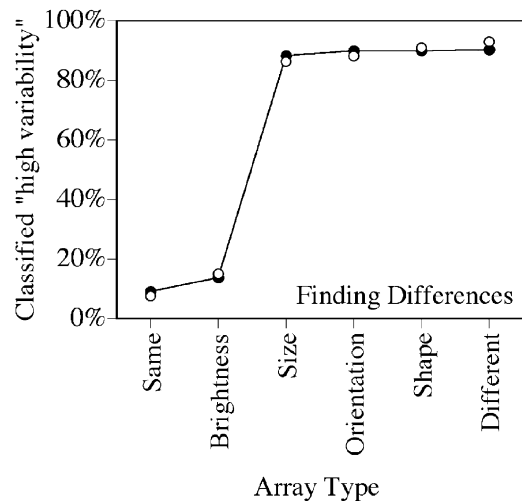


Fig. 9. The simulation results for the Similarity study. The X-axis indicates both the Same and Different arrays and the arrays in which one of the dimensions was held constant. The plots show the best fitting Finding Differences Model (solid symbols connected by lines) and the actual outcome from Young et al. (2002a) (open symbols). The optimal m and b parameters for the model were 0.141 and 17.50 , respectively.

in the first set of simulations also generalized well to very different testing situations. We did observe variations in the best fitting m and b parameters across data sets; but, these differences may merely reflect variations across studies in the overall level of discriminative performance achieved before testing.

The only clear difference between the predictive power of the models was found in fitting the Similarity data set. Only the Finding Differences Model could incorporate the effects of degree of similarity and the differential salience of each dimension. Thus, the new model (with one less parameter) provides fits that are generally as good as those provided by the Positional Entropy Model and offers something more—the ability to fit a new data set that was problematic for the original model.

3. Generality of the model

The Finding Differences Model better captures the stimulus features that control behavior in our task—the number and degree of localized differences in a collection of items. But, does this finding generalize across species? Perhaps pigeons respond to aggregated local differences, but baboons and humans respond to positional entropy or to some other property of these collections. Or, perhaps pigeons, baboons, and humans differ quantitatively, but not qualitatively, in their performance on these tasks. Work underway will help us to answer those questions.

In a related task, Cook and co-workers have examined the pigeon's ability to discriminate displays that contain odd items or odd areas from displays that do not (e.g. Cook et al., 1995; Cook et al., 1997). Their pigeons must identify the location of differences. The activation map that underlies our Finding Differences Model may provide exactly the information that a pigeon needs to solve this task. Although pigeons in our variability discrimination tasks can ignore the location of differences, pigeons in the oddity detection tasks may need to leverage the location information that is present in the activation maps. The potential power of the Finding Differences Model to identify the stimulus properties that control behavior in variability discrimination, oddity detection, and visual

search tasks presents a possibility for parsimony that cannot be overlooked. The current model might also be extended to address the effects of temporal distance (reported in Young et al., 1997a, 1999) as an analog to spatial distance.

4. Final thoughts

Ours is not the first model that has tried to account for the effects of mixture, organization, and similarity in a variability discrimination task. Hoch et al. (1999) examined people's judgments of product variety on store shelves and offered a model that bears a family resemblance to the Finding Differences Model. Their approach was quite different from ours, preferring a model with many more free parameters to empirically determine, (a) the shape of the generalization function, (b) the class of functions that might model the effect of spatial distance, and (c) the utility of assuming uniform or varied saliences across dimensions. Given the limited data from their single study, they could only make broad conclusions concerning these functional relationships. Our preference was to make many of our modeling decisions based on data from other cognitive tasks, thus opting for a city-block distance metric, exponential effects of spatial distance, and varied dimensional salience. Be that as it may, Hoch et al.'s results and the success of their model provide a comfortable validation of our own conclusion: namely, that variability discrimination is all about finding differences.

It may be highly adaptive for organisms to find differences, whether in the service of a visual search task (e.g. finding the tiger in the grass) or a variability discrimination task (e.g. identifying which flock has the greater variety in potential mates). Furthermore, we effortlessly find differences when we shop (assessing product variety), look at an audience (assessing racial or age diversity), and visit an art museum (assessing the represented variations in styles, artists, and exhibits). Understanding this discrimination process will undoubtedly affect how marketers and museum curators organize their wares in order to manipulate our perceptions. Given that “variety is the very spice of life,” perhaps such efforts are in our own best interests, even if some of the variety is illusory.

Acknowledgements

This research was supported by National Science Foundation Grant 9904569.

References

- Baumol, W.J., Ide, E.A., 1956. Variety in retailing. *Manage. Sci.* 3, 93–101.
- Berlyne, D.E., 1960. *Conflict, Arousal, and Curiosity*. McGraw-Hill, New York.
- Broniarczyk, S.M., Hoyer, W.D., McAlister, L., 1998. Consumers' perceptions of the assortment offered in a grocery category: the impact of item reduction. *J. Mark. Res.* 35, 166–176.
- Cook, R.G., Cavoto, K.K., Cavoto, B.R., 1995. Same-different texture discrimination and concept learning by pigeons. *J. Exp. Psych.: Anim. Behav. Process* 21, 253–260.
- Cook, R.G., Katz, J.S., Cavoto, B.R., 1997. Pigeon same-different concept learning with multiple stimulus classes. *J. Exp. Psych.: Anim. Behav. Process* 23, 417–433.
- Duncan, J., Humphreys, G.W., 1989. Visual search and stimulus similarity. *Psych. Rev.* 96, 433–458.
- Hoch, S.J., Bradlow, E.T., Wansink, B., 1999. The variety of an assortment. *Mark. Sci.* 18, 527–546.
- Munsinger, H., 1966. Multivariate analysis of preference for variability. *J. Exp. Psych.* 71, 889–895.
- Munsinger, H., Kessen, W., 1966. Preference and recall of stimulus variability. *J. Exp. Psych.* 72, 311–312.
- Myung, I.J., 2000. The importance of complexity in model selection. *J. Math. Psych.* 44, 190–204.
- Pitt, M.A., Myung, I.J., Zhang, S., 2002. Toward a method of selecting among computational models of cognition. *Psych. Rev.* 109, 472–491.
- Shannon, C.E., Weaver, W., 1949. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana.
- Shepard, R.N., 1964. Attention and the metric structure of the stimulus space. *J. Math. Psych.* 1, 54–87.
- Shepard, R.N., 1991. Integrality versus separability of stimulus dimensions: from an early convergence of evidence to a proposed theoretical basis. In: Lockhead, G.R., Pomerantz, J.R. (Eds.), *The Perception of Structure: Essays in Honor of Wendell R. Garner*. American Psychological Association, Washington, DC, pp. 53–71.
- Treisman, A.M., Gelade, G., 1980. A feature-integration theory of attention. *Cogn. Psych.* 12, 97–136.
- Wasserman, E.A., Hugart, J.A., Kirkpatrick-Steger, K., 1995. Pigeons show same-different conceptualization after training with complex visual stimuli. *J. Exp. Psych.: Anim. Behav. Process.* 21, 248–252.
- Wasserman, E.A., Young, M.E., Nolan, B.C., 2000. Display variability and spatial organization as contributors to the pigeon's discrimination of complex visual stimuli. *J. Exp. Psych.: Anim. Behav. Process.* 26, 133–143.
- Wasserman, E.A., Fagot, J., Young, M.E., 2001. Same-different conceptualization by baboons (*Papio papio*). *J. Comp. Psych.* 115, 42–52.
- Wolfe, J.M., 1994. Guided search 2.0: a revised model of visual search. *Psychonom. Bull. Rev.* 1, 202–238.
- Young, M.E., Ellefson, M.R., 2002. Multidimensional variability discrimination.
- Young, M.E., Ellefson, M.R., in press. The joint contributions of shape and color to variability discrimination. *Learn. Motiv.* 34, 52–67.
- Young, M.E., Wasserman, E.A., 1997. Entropy detection by pigeons: response to mixed visual displays after same-different discrimination training. *J. Exp. Psych.: Anim. Behav. Process.* 23, 157–170.
- Young, M.E., Wasserman, E.A., 2001. Entropy and variability discrimination. *J. Exp. Psych.: Learn. Mem. Cogn.* 27, 278–293.
- Young, M.E., Wasserman, E.A., 2002. Detecting variety: what's so special about sameness? *Journal of experimental psychology: General* 131, 131–143.
- Young, M.E., Wasserman, E.A., Dalrymple, R.M., 1997a. Memory-based same-different conceptualization by pigeons. *Psychonom. Bull. Rev.* 4, 552–558.
- Young, M.E., Wasserman, E.A., Garner, K.L., 1997b. Effects of number of items on the pigeon's discrimination of same from different visual displays. *J. Exp. Psych.: Anim. Behav. Process.* 23, 491–501.
- Young, M.E., Wasserman, E.A., Hilfers, M.A., Dalrymple, R.M., 1999. An examination of the pigeon's variability discrimination using lists of successively presented stimuli. *J. Exp. Psych.: Anim. Behav. Process.* 25, 475–490.
- Young, M.E., Castro, L., Wasserman, E.A., 2002a. Multidimensional variability discrimination in the pigeon.
- Young, M.E., Ellefson, M.R., Wasserman, E.A., 2002b. The Finding Differences Model of variability discrimination.