

Perceptual Sampling of Orthogonal Straight Line Features*

James T. Townsend**, Gary G. Hu, and F. Gregory Ashby

Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907, USA

Summary. Among mathematical models of visual confusion of letters and similar material, those that posit a feature detection process have been especially popular. The present study provides direct tests of several of the central assumptions of such models with feature-stimuli composed of the blank, one of two straight line features, or both line features positioned at a right angle. In one condition, the two features were connected when they appeared together, whereas in the other condition they were separated by a gap. A model which makes the strong assumptions that the features are sampled ('detected') independently and then reported in a direct, unbiased fashion, performed acceptably in both conditions. Feature dependency models and those positing a biased decision process were ruled out on the basis of poor fits or lack of parsimony. The perceptibility (d') of a specific feature depended on the stimulus that contained it in the Gap condition but not in the Connected condition. The relative perceptibility of the horizontal vs in the vertical features was also different in the Gap vs Connected conditions. The results were compared with other recent studies, including ones in which sampling independence was falsified, apparently because of greater stimulus complexity, and employing a stimulus set that did not contain all possible combinations of features.

*We would like to thank J.T. Berning for help in conducting the experiments. The first and second authors were supported by NSF Grant #BNS 76–84053 and the third author by Purdue Research Foundation Grant #XR 0104 during the early portions of this research. The latter portions were aided by National Science Foundation Grant #BNS 79–20298 to the first author. The third author is now at Human Performance Center, Ohio State University, Columbus, Ohio 43210. Send for reprints to Dr. James T. Townsend, Department of Psychological Sciences, Purdue University, West Lafayette, Indiana 47907, USA

**Corresponding author

Introduction

The concept of mental feature representations has played an important role over a wide spectrum of areas in cognitive psychology in recent years. However, it has seemed particularly viable in visual perception, no doubt owing in large part to the noteworthy physiological results on various species by Hartline (1940); Lettvin, Maturana, and Pitts (1959); Hubel and Wiesel (1962); and, lately, many others. These findings indicated the existence of neurons in the visual system in various species which responded more or less exclusively to the presence of specific aspects of visual displays. In particular, Hubel and Wiesel were primarily responsible for demonstrating the operation of what appear to be elementary geometric feature detection neurons in visual cortical areas 17, 18, and 19.

Although contemporary feature theories are not capable for explaining all aspects of pattern perception, they are nevertheless highly attractive in that they offer a powerful alternative to strict template matching (see, e.g., Reed 1973; Massaro and Schmuller 1975); especially in the recognition of simple forms such as letters. Provided one is prepared to hypothesize which parts of a set of stimuli constitute the features, it is then possible to develop and test mathematical models of pattern recognition based on feature detection in psychophysical settings (e.g., Rumelhart 1971; Geyer and DeWald 1973). More recently, Wandmacher (1976) and Townsend and Ashby (1976) have begun direct investigation of the axioms on which models are based. The present study continues this investigation with particular emphasis on whether or not basic features are sampled independently of one another.¹

A popular method of testing hypotheses and models concerning the recognition process is to compare them with an experimentally generated confusion matrix. A confusion matrix has as rows the set of stimuli and as columns, the set of corresponding responses. In the situation with which we are concerned, each stimulus is associated with exactly one response and $P(R_j|S_i)$ is the probability that response j is given after presentation of stimulus i . That quantity is thus the (i,j) entry in a confusion matrix. Similarly, $\hat{P}(R_j|S_i)$ is the corresponding estimated probability. That is, $\hat{P}(R_j|S_i)$ is the observed proportion of times that response j was given to stimulus i in an experiment; it may be compared with the theoretically predicted $P(R_j|S_i)$ after the parameters of a model are given numerical values. The indices i and j will run from 1 to N where N is the number of stimuli and responses. $P(R_i|S_i)$ is thus the probability correct given stimulus i and is placed in the (i,i) diagonal position of the confusion matrix.

Mathematically specified models of perceptual confusion have generally been composed of two basic parts, a sensory part and a decision part. The sensory part typically contains structure and parameters supposed to represent sensitivity, signal strength, and stimulus similarity factors, whereas the decision part contains structure and parameters that represent response bias, motivational and learning factors. Although decisional structure will play a role in our analyses, the sensory portion of the process will be our primary focus in this paper.

The sensory assumptions of previous feature models that have actually been applied to confusion data (sometimes implicitly) have been one or more of those listed below.

¹ Because the terms 'feature extraction' and 'detection' often carry connotations of questionable peripheral physiology, we prefer the more neutral term 'feature sampling' in general psychophysical settings.

Feature Sampling Assumptions

1. *Presence or Absence of a Feature in a Stimulus.* A certain feature either is contained in a stimulus in full-blown form or is entirely absent.
2. *All-or-none Feature Sampling.* Features are detected or not. There are no in-between representations, although it is not ruled out that an observation on a sensory continuum could occur after which a yes-no decision is made as to the presence or absence of the feature.
3. *High Feature Discriminability of Sampling.* Features are not misperceived as one another. That is, a unique feature is never 'sampled' as a different feature.
4. *Sampling Independence.* This assumption states that the sampling (i.e., extraction, detection) of a feature is probabilistically independent of the sampling of any other. Thus, the probability that any particular subset of features is sampled can be written as the product of the separate probabilities of the individual features.
5. *Across-stimulus Invariance.* This assumption means that the average probability of a particular feature being sampled does not depend on the particular stimulus of which it is a part. One consequence of this assumption is that feature sampling probabilities do not depend on the number of features contained in a stimulus and that the capacity at this level must therefore be unlimited.
6. *Across-feature Invariance.* Here it is supposed that all features within a given stimulus possess the same average probability of being sampled.
7. *High Threshold Feature Sampling.* Features may be lost from a presented stimulus, but not gained (sampled) when they are not present in a stimulus. Thus, viewed in the context of signal detection theory, feature 'misses' can occur but not feature 'false alarms'. A feature present in the stimulus will be referred to as a 'real feature', whereas a feature that might be reported even though it is not contained in the stimulus will be called a 'ghost feature'.

After some subset of features has been sampled in the sensory process, it is typically assumed that this sample is matched in some fashion against the individual sets of features that make up the various possible stimuli. A subset of potential response alternatives is formed (often called the 'confusion' or 'candidate' set; e.g., Geyer and DeWald 1973; Rumelhart 1971), the particular members of which depend on the degree of similarity each bears to the sampled feature set. This matching and confusion-set-formation process occupies a stage intermediate to the sensory phase and the final decision mechanism, but may be included in the decision phase for convenience.

Finally, it is supposed that in the final decision phase, a response is selected from the confusion set according to a set of probabilities on those alternatives belonging to the confusion set. Various assumptions may be made about these probabilities, but the most parsimonious, and at the same time presently viable, is to postulate that each response is associated with a response strength (represented by a positive number) and that the probability of a particular response from the confusion set is given by the ratio of that particular response strength to the sum of all the response strengths of the alternatives in the confusion set. Note that this tacitly assumes that the response strength is invariant across

the various confusion sets, an assumption recently questioned (Lappin 1978). However, it is not easy to come up with a viable replacement that does not generate an unmanageable number of parameters. The possibility that Bayesian analyses govern this stage does not appear to be justified (Wandmacher 1976).

Now let us discuss the sensory assumptions in a little more detail. Assumptions (1), (2), and (3) occur in all of the feature models fit to visual confusion data and are accepted as postulates here. With regard to (1), Garner (1978) has defined a 'feature' as either existing or not in a stimulus, without any in-between levels. Whether or not one agrees with this definition in any ultimate sense, all well-specified mathematical models of confusion make this assumption. In some 'realistic' settings it might be necessary to assume that features can exist on a continuum, but in many situations the present assumption should suffice. Assumption (2) is probably reasonable as a first approximation, particularly since the all-or-none 'threshold' may be at the decisional level. However, situations can definitely be created where an explicit sensory continuum should be postulated (see, e.g., Oden 1979). The third assumption almost certainly holds when the features are very distinct, as is the case in most of the previous experiments where explicit modeling of the feature processes occurred.

Assumption (4) is in some ways the most critical because if sampling independence fails then one must know the rules of dependence in order to formulate a testable model. Alternatively, feature similarity parameters might be written directly into $P(R_j | S_1)$ formulae without specification of an underlying process. Such models are less attractive than models based on a more fully specified set of mechanisms.

Number (5) would hold if the perceptibility of a feature is independent of the rest of the stimulus in which it is contained. Assumption (6) would be in force if all features in a given stimulus were equally perceptible. Finally, (7) has generally been postulated in mathematical feature models, in spite of the influence that the theory of signal detectability has had in arguing against so-called high threshold notions (e.g., Green and Swets 1966).

A brief presentation of the most critical previous investigations is in order.² In a study noteworthy for its experimental and theoretical conception, Wandmacher (1976) investigated a class of models which included feature sampling independent as well as dependent alternatives. Wandmacher's six stimuli were composed of one or two straight lines, the latter being affixed at an acute angle. Sampling independence was supported (Wandmacher's stochastic independence), across-stimulus invariance (Wandmacher's context independence) was not, nor was across-feature invariance (his 'equal probability' constraint). The high threshold assumption (7) could not be tested directly because no blank stimuli were included in the displays.

Townsend and Ashby (1976) employed four block letter stimuli constructed out of equal length straight lines connected at right angles. The observers (henceforth referred to as *O*s) were required to report the features they thought they saw as well as to make

² Another type of paradigm has been employed in which a presented feature is one of two values on some dimension, rather than, say, the presence of a particular line or curve. For example, Schulze, Baurichter, Gerling, and Grobe (1977) employed two angle sizes as one of their binary 'features.' The state of knowledge in pattern perception is not yet sufficiently advanced to permit firm theoretical linkages between this and the present type of design.

a letter decision. This permitted a tentative separation of the sensory and decision phases of the hypothetical underlying process. Strong sampling dependencies were found in the feature reports and the overall confusion matrices were as well predicted by a much less detailed non-feature model (the choice model, Luce 1963; Townsend 1971, a, b). However, across-stimulus invariance held acceptably, and there was some evidence that the degree of across-feature variability that existed was due to criterial influences rather than purely sensory effects. Further, the stimuli were quite homogeneous, having about the same number of features (three or four) connected at right angles, so that may be the reason that across-stimulus invariance was found in that study but not in Wandmacher's. That is, there may be a greater difference in the perceptibility of a line by itself as compared with another line at an acute angle than there is between three and four lines connected at right angles. Finally, the high threshold assumption was falsified.

It was not clear why Wandmacher found sampling independence while Townsend and Ashby found dependence. One hypothesis is that independence can appear only in the most simple of stimulus situations. Another is that certain subsets of sampled features seem to the *O* as being impossible and are thus disregarded, or otherwise perturbing the feature reports. This latter alternative will be considered briefly in the discussion section.

The former possibility was addressed in initial analyses of a confusion study in which stimuli were composed of one or two straight lines (Townsend, Hu and Ashby 1980). The two-lined stimuli were set at right angles to one another and were either physically connected or separated by a gap (hereafter referred to as the *Connected and Gap* conditions). Both lines, neither line (the blank stimulus), or either line separately could be the stimulus in any given block of trials.

Two experiments were performed. The first used three separate feature conditions: Γ , \vdash , \perp as constituting the component features employed in a block of trials. The second experiment used only Γ . In addition, the first experiment included Gap and Connected trials within blocks, whereas the second experiment tested Gap vs Connection perception in separate blocks. These Gap and Connected conditions were meant to ascertain whether sampling independence might be stronger with a greater distance between features. Thus, a lateral interference hypothesis might predict that the connected features would evidence a negative dependence, whereas the separated features might show a diminished or zero dependence. A preliminary report (Townsend, Hu, and Ashby 1980) presented analyses of sampling independence on the set of trials when both lines were present. Experiment 1 was flawed by an artifact which was corrected in Experiment 2, but basically the results provided reasonable support for independence.

The present paper extends the investigation to explanation of the entire confusion matrix; that is, to an examination of the response frequencies based on all the four types of stimulus presentation (neither, either, or both lines present in the stimulus). Because of the artifact present in the first experiment, the following treatment will be confined to Experiment 2. Suffice it to say that with models attuned to handle the artifact, the basic conclusions are in agreement with those of Experiment 2, albeit with model fits not quite so accurate. Henceforth, all references will be to Experiment 2. The models to be tested are given next.

The Models

In the preliminary analyses mentioned above (Townsend, Hu and Ashby 1980), a sampling independence model which assumed that features were reported if and only if they were sampled, provided reasonable fits to the analyzed data. That is, it was not necessary to assume that any alternative other than what was sampled was let into the candidate set from which the response was made. Thus, if l was sampled, then l was the response made, and so on.

In fitting the entire confusion matrix, it was found that this same assumption of direct report again provided reasonably good fits to the data and thus it forms the basis of the models we considered below. It should be mentioned that after the direct report models were analyzed a great many other models which included a biased decision stage were tested. None performed nearly as well as the models developed below, so they have been excluded from consideration. (However, a particular example will be mentioned later.) This strong assumption of direct report formed the basis of the primary model tested against the entire data set which, for obvious reasons, we call the 'independent direct report model' (henceforth referred to as the IDR model).

More formally the IDR model is characterized by assumptions (1), (2), (3), and (4). Whether or not assumption (5) holds apparently depends on the specific stimuli employed so models that do or do not assume that a feature has the same sampling probability in different stimuli will be analyzed. No previous results with which we are familiar support assumption (6) so that different features may have distinct sampling probabilities. Assumption (7) is tested in our analyses.

Further, the IDR model thus generated will be tested against a sampling *dependent* model, hereafter called the DDR (dependent direct report) model. Because this model has quite a few parameters, it was necessary to require assumption (5), across-stimulus invariance. It is important to note that the present IDR model assumes that the sampling of ghost features is also independent, as did the models of Townsend and Ashby (1976). Similarly, the DDR model assumes dependence of both ghost and real features.

The next five definitions give explicit form to the tested models.

Definition 1 (Stimulus)

Define the stimulus presented as $S_V S_H$ where S_V and S_H are variables of the form

$$S_V = \begin{cases} V, & \text{vertical present} \\ \bar{V}, & \text{vertical absent} \end{cases} \quad S_H = \begin{cases} H, & \text{horizontal present} \\ \bar{H}, & \text{horizontal absent} \end{cases}$$

A line above S_V or S_H indicates the opposite 'value', where 'value' is defined by presence or absence of the feature. Thus, $S_V S_H = \bar{V} H$ indicates that the vertical line was not present in the stimulus but the horizontal line was.

Definition 2 (Sample)

Define the sampled set of features as $s_V s_H$ where s_V and s_H are variables of the form

$$s_V = \begin{cases} v, & \text{vertical present} \\ \bar{v}, & \text{vertical absent} \end{cases} \quad s_H = \begin{cases} h, & \text{horizontal present} \\ \bar{h}, & \text{horizontal absent} \end{cases}$$

A line above s_V or s_H indicates the oposite 'value', where 'value' is defined as presence or absence.

As an example, $s_V s_H = vh$ indicates both a vertical as well as a horizontal line were sampled, without reference to whether or not they were actually contained in the stimulus.

The next definition gives the general form of the sampling probability.

Definition 3 (Sampling Probability)

The probability that $s_V s_H$ is sampled given the presentation of stimulus $S_V S_H$ is written $P(s_V s_H | S_V S_H)$.

An example for this definition would be $P(v\bar{h} | V\bar{H})$ which is the probability that the vertical line was correctly sampled and the horizontal line was correctly not sampled.

We are now in a position to define the two major models investigated in this study.

Definition 4 (IDR Model: Assumes sampling independence)

- a. With Across-Stimulus Variability

$$P(s_V s_H | S_V S_H) = p(s_V | S_V S_H) \times p(s_H | S_V S_H)$$

- b. With Across-Stimulus Invariance

$$P(s_V s_H | S_V S_H) = p(s_V | S_V) \times p(s_H | S_H)$$

That is,

$$p(s_V | S_V S_H) = p(s_V | S_V \bar{S}_H) = p(s_V | S_V) \text{ and } p(s_H | S_V S_H) = p(s_H | \bar{S}_V S_H) = p(s_H | S_H).$$

The interpretation of part (b) of Def. 4 is that the probability of sampling the vertical or horizontal line, whether perceived as a ghost or real feature, is the same whether the other feature is or is not present.

In line with the IDR model with across-stimulus variability of sampling probabilities, the example below Def. 3 would be expressed as

$$P(v\bar{h} | V\bar{H}) = p(v | V\bar{H}) \times p(\bar{h} | V\bar{H}) = p(v | V\bar{H}) \times [1 - p(h | V\bar{H})].$$

The differentiating factor between the previous model type and the *dependence* direct report model (DDR model) is that the joint probabilities of feature sampling must be incorporated into the formulae. This can be done using one conditional and one marginal probability according to the well-known formula $P(A \& B) = P(A) \times P(B|A)$, as the following definition will indicate. Thus, let A play the role of s_H and B the role of s_V . Conditioning on the stimuli will remain the same.

Definition 5 (DDR Model: Assumes sampling dependence)

- a. With Across-Stimulus Variability

$$P(s_V s_H | S_V S_H) = p(s_H | S_V S_H) \times p(s_V | s_H, S_V S_H)$$

- b. With Across-Stimulus Invariance

$$P(s_V s_H | S_V S_H) = p(s_H | S_H) \times p(s_V | s_H, S_V)$$

$$p(s_V | s_H, S_V S_H) = p(s_V | s_H, S_V \bar{S}_H) = p(s_V | s_H, S_V) \text{ and } p(s_H | S_V S_H) = p(s_H | \bar{S}_V S_H) = p(s_H | S_H), \text{ in the case of across-stimulus invariance.}$$

Clearly, a positive dependency implies $p(s_V|s_H, S_V S_H) > p(s_V|\bar{s}_H, S_V S_H)$, a negative dependency the reverse inequality and independence implies equality.

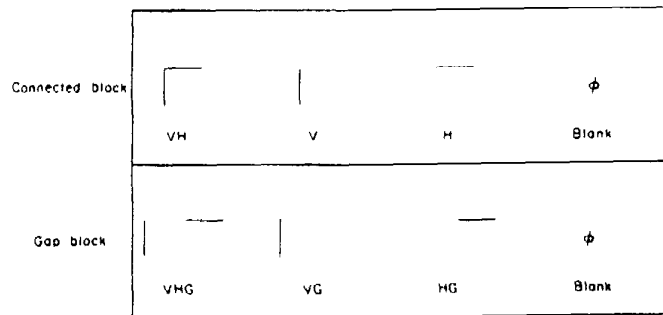
Experimental Method

Subjects

The *O*s were four Purdue upper division majors in psychology with 20/20 vision.

Apparatus and Stimuli

A Gerbrands two-field tachistoscope (Model T-2b) was used to present the stimuli. The operational stimulus features are shown in Figure 1 according to which type of block they were presented in. G indicates a gap between the two lines in the stimulus and \emptyset is the symbol used to denote the blank display. Whenever the stimulus feature V was presented in the gap condition, it was displaced approximately 12.5 min to the left of its position in the Connected blocks. Similarly, stimulus feature H was displaced 12.5 min distance to the right in all Gap blocks. When both were presented, they were, of course, separated by a 'gap' angle of 25 min. The length of each line also subtended an angle of 25 min.



A prestimulus fixation field was described by a set of four dots which were arranged as the corners of a square with the stimulus in the center. The four dots were on the screen at all times except during the brief intervals of stimulus presentation. The fixation field on any one side subtended an angle of about 2° at *O*s eye. Luminance was maintained at 27.4 cd/m^2 . Responses were given verbally and were recorded by the E on recording sheets.

Procedure

The Gap and Connected conditions were presented in separate counterbalanced blocks, generating 4×4 confusion matrices for each *O* in each condition. The *O*s had five days of practice and calibration for a total of five hours. Before the experiment proper began, stimulus duration was set for each *O* individually so that the proportion correct was approximately 0.45. Twenty trials of practice preceded each of fifteen experimental sessions. Each of the four stimuli was presented 10 times in each block of trials. There were thus 40 trials per block and because each block was run once a day, there were 80 experimental trials per session producing a total of 150 trials per stimulus per *O* over the course of the experiment.

That is, each row in each of the two confusion matrices was based on 150 trials, for each O . The stimuli were randomized across trials within the foregoing constraints. O s were instructed to report only the features they thought they had seen, to attempt to disregard and not to report any gaps.

Results: Basic Model Tests and Perceptibility Analyses

We are interested in the individual processes, of course, so Tab. 1 shows the 4×4 confusion matrices for each O for the Gap and Connected conditions as well as group averages. The theoretical entries will be explained shortly.

It can be seen that there was a tendency for accuracy (represented by diagonal entries) to be somewhat greater for VH p(vh|VH) in the Connected condition over the Gap condition, but the reverse was true in the case of V. The H and \emptyset stimuli did not evidence a strong difference in accuracy in the two spacing conditions. Further, the estimated probability correct, averaged across stimuli and O s was the same for the Connected and Gap conditions (Ave $\hat{P}(\text{Cor}) = 0.54$). Also, the H stimulus was always more accurately reported than the V stimulus.

The Gap and Connected conditions were analyzed separately. The IDR and DDR models were fit to the individual confusion matrices with one version of each assuming across-stimulus invariance and one version assuming across-stimulus variability of feature sampling probabilities.

We can, of course, estimate the models' parameters and simultaneously test the models using the now classic Chandler (1965) χ^2 minimization program. Moreover, if one model is a nested case of another, then the difference of the two fitted χ^2 's is again approximately χ^2 distributed with *df* given by the difference in *dfs* between the less and more general models. Thus, assuming across-stimulus invariance, the IDR model is a nested case of DDR. Similarly, assuming independence, the across-stimulus invariant model is a nested case of the across-stimulus variability model. The nested tests reported below were all carried out at the individual O level.

Within the Connected condition, the IDR model with across-stimulus variability was not significantly better than IDR with across-stimulus invariance for any of the O s. (According to the nested χ^2 test given above, $\chi^2_{0.05}(\text{df}) = \chi^2_{0.05}(4) = 9.5$, where *df* = 4 is the difference in degrees of freedom in the two models.) Therefore, the IDR version with across-stimulus invariance is the preferred model. Next, this model is also a nested case of the DDR model with across-stimulus invariance, so may be statistically tested against it. Three O s evidenced no significant difference at 0.05 level ($\chi^2_{0.05}(2) = 6.0$), but O_1 showed significance at about the 0.01 level. In addition, in the original χ^2 tests of the models, the IDR model with across-stimulus invariance fit all of the O s reasonably well at the 0.05 level (O_4 was marginally significant with a $\chi^2 = 15.7$ vs the significance criterion $\chi^2_{0.05}(8) = 15.5$).

The best fitting IDR model in the Gap condition required across-stimulus variability (only one O did not exhibit a significant difference ($\chi^2_{0.05}(4) = 9.5$) in the nested test). Because the DDR model could only be employed with across-stimulus invariance, a nested test against across-stimulus variability is not permitted. In the separate model fits in this condition, each model (IDR and DDR) failed to fit one of the O s at the 0.05 level (not the same O). At a qualitative level, it appeared that the IDR and DDR models fitted one O somewhat better than the other model did, with the other two O s being fitted about equal-

Table 1. Observed and predicted confusion matrices

Response:	VH		V		H		ϕ	
	Gap	Obs	Gap	Obs	Gap	Obs	Gap	Obs
<i>Observer:</i>	Obs	PRE ^a	Obs	PRE ^a	Obs	PRE ^a	Obs	PRE ^a
	Obs	PRE ^b	Obs	PRE ^b	Obs	PRE ^b	Obs	PRE ^b
O ₁	.16	.19	.20	.18	.35	.32	.38	.31
O ₂	.52	.53	.57	.20	.22	.20	.29	.07
O ₃	.26	.28	.33	.26	.27	.24	.27	.22
O ₄	.19	.19	.26	.16	.35	.34	.37	.30
AVE	.28	.30	.34	.20	.29	.28	.33	.22
O ₁	.07	.06	.05	.47	.04	.05	.06	.41
O ₂	.08	.07	.05	.58	.57	.04	.05	.31
O ₃	.03	.03	.03	.57	.03	.02	.03	.37
O ₄	.07	.09	.07	.42	.43	.09	.08	.41
AVE	.06	.06	.05	.51	.51	.05	.05	.38
O ₁	.03	.03	.09	.02	.02	.50	.45	.44
O ₂	.07	.08	.10	.03	.03	.64	.72	.25
O ₃	.01	.02	.03	.01	.02	.60	.58	.36
O ₄	.02	.04	.01	.03	.04	.55	.54	.38
AVE	.04	.04	.06	.02	.03	.57	.57	.36
O ₁	.03	.01	.04	.08	.03	.08	.08	.83
O ₂	.05	.03	.03	.14	.10	.14	.14	.69
O ₃	.02	.01	.01	.09	.08	.09	.04	.80
O ₄	.03	.02	.03	.07	.11	.20	.09	.70
AVE	.03	.02	.03	.10	.07	.13	.09	.76

^aPrediction from IDR with across-symbol variability and allowing ghost features

^bPrediction from IDR with across-symbol invariance and allowing ghost features. See text for further explanation

ly well by either model. Thus, one cannot decisively falsify the DDR model in this condition. However, overall, the most parsimonious model for the entire corpus of data would be the IDR model, invested with across-stimulus invariance to predict the Connected data, but requiring distinct sampling probabilities with different stimuli in the Gap condition. The predictions of the IDR model with across-stimulus invariance in the Connected condition and across-stimulus variability for the Gap condition are shown in Table 1. Its parameter estimates and χ^2 values are given in Table 2. As can be seen in those tables, the fits appear to be in good agreement with the data. There seems to have been a discernible dependence for one O in the Connected condition and a different O in the Gap condition, although statistical sampling error may have played a role there.

The important question, however, is not whether *no* dependencies at all exist in the sampling process, but rather: 1) whether independence provides a reasonable first approximation to the data, and 2) how it continues to hold or fails to hold under varying stimulus and contextual conditions. The first seems to be true in the present experiment. With regard to the second question, we need to draw on other experimental results, which we will do below. But first let us briefly consider some potential conclusions to be drawn from Tables 1 and 2.

From Table 1 it can be seen that there was a slight tendency for the models to overpredict probability correct, the only observer-consistent exception being the VH stimulus in the Gap condition. Whether these minor effects are due to some real second-order difficulties in the models or only to statistical error is not known at the present.

From Table 2 it is apparent that in the Gap condition, sampling probabilities tend to be higher in some conditions than others. However, the estimated 'real' and 'ghost' feature sampling probabilities correspond respectively to hit and false alarm frequencies. Therefore, according to our theoretical orientation, these may be used to calculate d' and β values. The assumption of independent channels for the two features makes these computations meaningful. Table 3 reports these, and it can be verified that in several cases the sampling probabilities are higher in one condition than another but with the respective d 's reversed.

Without verbally tracing all the pertinent comparisons in Table 2, the major finding with respect to feature perceptibility as reflected in d' are as follows. In the Connected condition, the horizontal feature was more perceptible than the vertical feature. In the Gap condition, perceptibility of the vertical feature was improved to the extent that when both features were present in the stimulus, it was seen even better than the horizontal feature. If only one feature was present in the stimulus, the vertical and horizontal features were perceived about equally well. However, in the Gap condition even the horizontal feature was perceived slightly better with the vertical feature present than when it was absent. These findings can be explained if attention was greater on the vertical feature in the Gap condition and on the horizontal feature in the Connected condition. In addition, there appears to have been a slight Gestalt influence operating in the Gap condition which improved performance when both features were present. We do not know at this point what could have caused such an effect in the Gap but not in the Connected condition. It is crucial to note, though, that such an effect does not imply the presence of sampling dependencies. Incidentally, note that the estimated sampling probabilities decisively falsify assumption (7), the high-threshold axiom.

Table 2. Parameter estimates and χ^2 values of the best independence models

	Real				Ghost				χ^2
	$\hat{p}(v/VH)$	$\hat{p}(v/\overline{VH})$	$\hat{p}(h/VH)$	$\hat{p}(h/\overline{VH})$	$\hat{p}(v/VH)$	$\hat{p}(v/\overline{VH})$	$\hat{p}(h/VH)$	$\hat{p}(h/\overline{VH})$	
O ₁	.367	.536	.512	.533	.054	.091	.119	.091	18.37
O ₂	.730	.651	.730	.720	.107	.172	.112	.166	9.48
GAP ASV ^a	.533	.608	.520	.626	.037	.104	.055	.104	3.74
O ₃	.353	.507	.533	.594	.067	.092	.175	.223	5.59
O ₄	.496	.576	.574	.618	.066	.115	.115	.146	$\chi^2_{.05}(4) = 9.49$
	Real		Ghost		χ^2				
	$\hat{p}(v/V)$	$\hat{p}(h/H)$	$\hat{p}(v/V)$	$\hat{p}(h/H)$					
O ₁	.336	.552	.120	.121	13.42				
O ₂	.634	.817	.115	.139	9.51				
CONN ASI ^b	.520	.599	.077	.058	7.32				
O ₃	.432	.588	.114	.142	15.66				
O ₄	.480	.639	.106	.115	$\chi^2_{.05}(8) = 15.51$				

^a Across-stimulus variability^b Across-stimulus invariance

Table 3. A d' and β analysis of the vertical and horizontal estimated feature sampling probabilities

Based on	Hit	False alarm	Hit	False alarm	Hit	False alarm	Hit	False alarm
	$\hat{p}(v/VH)$	$\hat{p}(v/\sqrt{VH})$	$\hat{p}(v/VH)$	$\hat{p}(v/\sqrt{VH})$	$\hat{p}(h/VH)$	$\hat{p}(h/\sqrt{VH})$	$\hat{p}(h/VH)$	$\hat{p}(h/\sqrt{VH})$
	<i>1a</i>	<i>1b</i>	<i>2a</i>	<i>2b</i>	<i>3a</i>	<i>3b</i>	<i>4a</i>	<i>4b</i>
	\hat{d}'_v	$\hat{\beta}_v$	\hat{d}'_v	$\hat{\beta}_v$	\hat{d}'_h	$\hat{\beta}_h$	\hat{d}'_h	$\hat{\beta}_h$
O ₁	1.27	3.45	1.42	2.44	1.21	2.00	1.42	2.45
O ₂	1.86	1.78	1.34	1.45	1.82	1.75	1.55	1.35
O ₃	1.92	5.42	1.54	2.13	1.65	3.59	1.58	2.10
O ₄	1.88	2.86	1.35	2.42	1.02	1.55	1.00	1.30
AVE	1.73	3.38	1.41	2.11	1.42	2.22	1.39	1.80
Based on	Hit	False alarm	Hit	False alarm	Hit	False alarm	Hit	False alarm
	$\hat{p}(v/V)$	$\hat{p}(v/\sqrt{V})$	$\hat{p}(h/H)$	$\hat{p}(h/\sqrt{H})$	$\hat{p}(h/H)$	$\hat{p}(h/\sqrt{H})$	$\hat{p}(h/H)$	$\hat{p}(h/\sqrt{H})$
	<i>1a</i>	<i>1b</i>	<i>2a</i>	<i>2b</i>	<i>3a</i>	<i>3b</i>	<i>4a</i>	<i>4b</i>
	\hat{d}'_v	$\hat{\beta}_v$	\hat{d}'_h	$\hat{\beta}_h$	\hat{d}'_h	$\hat{\beta}_h$	\hat{d}'_h	$\hat{\beta}_h$
O ₁	.75	1.84	1.30	1.96	1.30	1.96	1.30	1.96
O ₂	1.54	1.94	1.99	1.19	1.99	1.19	1.99	1.19
O ₃	1.48	2.74	1.82	3.32	1.82	3.32	1.82	3.32
O ₄	1.03	2.02	1.29	1.73	1.29	1.73	1.29	1.73
AVE	1.20	2.14	1.60	2.05	1.60	2.05	1.60	2.05

^a Across-stimulus variability

^b Across-stimulus invariance

Discussion

Our data analyses definitely rejected the high threshold postulate (e.g., note the ghost feature sampling probabilities in Table 2). Further, we found that the assumption of a guessing process in conjunction with this independence model could be soundly rejected. In fact, a general model which did allow ghost feature sampling *as well as* an ensuing biased decision process (not mentioned previously) performed extremely poorly ($\chi^2 \cong 155$). It is impossible for the reduction in parameters caused by imposing a high threshold assumption on the foregoing general model to reduce the χ^2 to a reasonable number. Thus, a low threshold model assuming direct report of the sampled features appears to be the preferred explanation for the present data.

On the other hand, the success of the high threshold models in Wandmacher's (1976) study is somewhat troubling. It is true that the processing by *O*s may be somewhat different depending on exact circumstances, even in experiments in which the stimuli are sufficiently simple to facilitate sampling independence. Nevertheless, it is difficult to conceive how ghost features could have been entirely eliminated in the Wandmacher (1976) design. If stimulus intensities were quite high, resulting in high accuracy, then *O*s could probably discriminate and therefore reject weak ghost features. However, the accuracies of Wandmacher's *O*s were set between 0.25 and 0.50 probability correct, a low to moderate accuracy level. Of course, the absence of blank trials makes it difficult to critically evaluate the high threshold assumption, especially when combined with a biased decision process subsequent to the sampling process.

Before a consideration of experiments that have produced strong dependencies, a very recent study by Wandmacher, Kammerer, and Glowalla (1980) should be mentioned. Across-stimulus invariance and sampling independence were studied with across-stimulus variability being rejected and sampling independence supported. The (high threshold) assumption that no ghost features occurred could not be tested. The stimuli in their second experiment contained from 2 to 4 connected features. Thus, the results are basically compatible with Wandmacher (1976) and the present work.

Of studies arguing against sampling independence, at least in a direct report sense, perhaps the earliest was Hubert (1972). However, only children were run as observers and only group averages were analyzed so the results are not really compatible to the other work discussed here.

The previously noted Townsend and Ashby (1976) study found dependencies that were quite larger. Table 4a exhibits the estimated joint probabilities and the independence predictions (the product of the estimated marginal probabilities) of report for a horizontal and vertical feature from that study. The features employed in Table 4a could not be used in Table 4b because they appeared in every letter used in the earlier study. However, Table 4b reproduces the estimated probabilities for an analogous feature-pair from the present study for purposes of comparison. Both tables also exhibit the differences between predicted and observed values as well as a Z-test on those differences. It is noteworthy that not only are the Townsend and Ashby differences very significant but, more importantly, are larger than those of the present study even though there were almost three times as many trials in the earlier work (resulting in a smaller standard error).

Table 4. Example comparison of independence test from Townsend and Ashby data (1976) and the present experiment

<i>a</i>					
Observer	Cond.	$\hat{P}(\bar{\cap})$	$p(\bar{\cap}) \times p(\bar{\cap})$	Diff.	Z-test (after arcsine) transformation
O ₁	FH	.212	.337	-.125	-4.08***
	AE	.350	.371	-.021	-.62
O ₂	FH	.102	.198	-.096	-3.91***
	AE	.091	.170	-.079	-3.47***
O ₃	FH	.212	.333	-.121	-3.91***
	AE	.355	.413	-.058	-1.74*

N = 420

* significant at $\alpha = .10$ (two-tailed)

*** significant at $\alpha = .001$ (two-tailed)

<i>b</i>					
O ₁	CONN	.189	.195	-.006	-.13
	GAP	.165	.201	-.036	-.81
O ₂	CONN	.560	.557	.003	.05
	GAP	.517	.536	-.019	-.33
O ₃	CONN	.333	.311	.022	.41
	GAP	.263	.281	-.018	-.35
O ₄	CONN	.260	.239	.021	.42
	GAP	.187	.189	-.002	-.04

N = 150

All Z-tests were nonsignificant at the $\alpha = .30$ level (two-tailed)

Note. The Townsend and Ashby data used in Table 4a are based on feature reports following stimuli which included both these features, as are the data from the present experiment

The Os in the Townsend and Ashby (1976) experiment were required to make a letter response as well as the feature report. Analyses revealed that they often responded with letter alternatives that might have led to the feature sample by losing features but virtually never with alternatives which would have had to gain features (i.e., even though they often reported ghost features, they tended to discount this possibility in their letter responses). This decision process could have led to a lack of independence if Os suppressed feature samples reflecting impossible (in *their* minds) combinations. Thus, if this were true, a combination like $_ _$ should rarely have been reported since no letters contained those features, and therefore could not have led to that sample by feature loss. This, in fact, occurred, so that factor accounts for at least some of the dependencies. However, this effect cannot predict the dependencies in such feature pairs as those in Table 4a. Nevertheless, an experimental paradigm that employed all feature combinations *plus* a blank feature in order to generate a more complex set of stimuli than Wandmacher (1976) or the present study used would certainly be of inter-

est, because no configurations would be impossible, either from a lost real feature or from a gained ghost feature point of view.

Townsend and Hu (1980) recently reported a confusion experiment meeting these conditions, based on non-letter stimuli composed of all possible combinations of three straight line features and one curved line, plus the blank stimulus. The stimuli were thus about as complex as real letters but without being overlearned as symbols. Very strong dependencies were found, interestingly, of a positive nature. That is, if one feature were sampled, then it was more likely that another would also be sampled. This result is contrary to what would be expected from lateral interference. Townsend and Ashby (1976) had found both positive as well as negative dependencies, possibly reflecting the decision strategy mentioned above as well as positive effects of the type found in the latter study.

It is of concern, however, that the particular stimuli of Wandmacher, Kammerer, and Glowalla (1980) which were multi-featured did not produce sampling dependencies. The reason for this is unclear. A model of the type that fitted their data performed very poorly with our data. Obviously, more work is in order with regard to ascertaining the reasons for the differences in findings as well as determining the cause of the positive sampling dependencies in the Townsend and Hu data. What seems to be emerging in a general way, however, is that feature sampling independence is an acceptable hypothesis in sufficiently restricted simple stimulus-confusion experiments, but may easily be lost in more complex situations. Finally, across-stimulus invariance of feature sampling probabilities appears to be viable only in a highly homogeneous set of stimuli.

References

- Chandler JP (1965) Subroutine STEPIT: An algorithm that finds the values of the parameters which minimize a given continuous function. Quantum Chemistry Program Exchange, Indiana University, Bloomington, IN
- Garner WR (1978) Aspects of a stimulus: Features, dimensions, and configurations. In: Rosch E, Lloyd BB (eds) *Cognition and Categorization*. L. Erlbaum Associates, Hillsdale, NJ
- Geyer LH, DeWald CG (1973) Feature lists and confusion matrices. *Percept Psychophys* 14:471–482
- Green DM, Swets JA (1966) *Signal Detection Theory and Psychophysics*. Wiley, New York, NY
- Hartline HK (1940) The receptive fields of optic nerve fibers. *Am J Physiol* 130:235–237
- Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *J Physiol (Lond)* 160:106–154
- Hubert LA (1972) A statistical method for investigating the perceptual confusions among geometric configurations. *J Math Psychol* 9:389–403
- Lappin JS (1978) The relativity of choice behavior and the effect of prior knowledge on the speed and accuracy of recognition. In: Castellan NJ, Jr, Restle F (eds) *Cognitive Theory*, vol 3. L. Erlbaum Associates, Hillsdale, NJ

- Lettvin JW, Maturana HR, McCulloch WS, Pitts WH (1959) What the frog's eye tells the frog's brain. *Proc IRE* 47:1940-1951
- Luce RD (1963) Detection and recognition. In: Luce RD, Bush RB, Galanter E (eds) *Handbook of Mathematical Psychology*, vol I. John Wiley & Sons, New York, NY
- Massaro DW, Schmuller J (1975) Visual features, preperceptual storage, and processing time in reading. In: *Understanding Language*. Academic Press, New York, NY
- Oden GC (1979) A fuzzy logical model of letter identification. *J Exp Psychol Hum Percept Perform* 5:336-352
- Reed SK (1973) *Psychological Processes in Pattern Recognition*. Academic Press, New York, NY
- Rumelhart DE (1971) A multicomponent theory of confusions among briefly exposed alphabetic characters. Center for Human Information Processing Report No 22. University of California, San Diego, CA
- Schulze HH, Baurichter W, Gerling W, Grobe R (1977) Independent feature processing in the visual system: The effects of symmetry and distance of components. *Psychol Res* 39 (3):249-259
- Townsend JT (1971a) Theoretical analysis of an alphabetic confusion matrix. *Percept Psychophys* 9:40-50
- Townsend JT (1971b) Alphabetic confusion: A test of models for individuals. *Percept Psychophys* 9:449-454
- Townsend JT, Ashby FG (1976) Testing contemporary models of letter recognition. Paper presented at the annual meetings of the Midwestern Psychological Association, Chicago, IL
- Townsend JT, Hu GG (1980) Testing the independence of visual feature perception. Paper presented at the Thirteenth Annual Mathematical Psychology Meetings, Madison, WI
- Townsend JT, Hu GG, Ashby FG (1980) A test of visual feature sampling independence with orthogonal straight lines. *Bull Psychonom Soc* 15 (3):163-166
- Wandmacher J (1976) Multicomponent theory of perception: Feature extraction and response decision in visual identification. *Psychol Res* 39:17-37
- Wandmacher J, Kammerer EM, Glowalla U (1980) Context dependence in visual feature processing. *Psychol Res* 40:335-351

Received May 26, 1981