Modeling feature perception in brief displays with evidence for positive interdependencies

J. T. TOWNSEND, G. G. HU, and R. J. EVANS
Purdue University, West Lafayette, Indiana

Recent studies have begun to test certain fundamental assumptions underlying popular models of feature perception. This research is continued here. It was discovered that every basic assumption that was open to test in the present study was disconfirmed. However, several new characteristics of line and curve feature perception were discovered: (1) Feature perception sensitivity was inversely related to the number of features present in a stimulus pattern, and the decision criteria for reporting a feature decreased with the number of features contained in a pattern; (2) the decrements in sensitivity reported in (1) were greater for features lying inside a pattern than for those on the exterior; and (3) feature perception sensitivity actually improved if another feature was known to be correctly perceived during the same trial. Likewise, feature sensitivity decreased if another feature was missed on any trial. At the present time, a system that first extracts global and then local (more detailed) featural information provides a basis that qualitatively accounts for our findings and is also compatible with several other studies in the literature.

One powerful method of studying pattern recognition is to present single members of a well-learned stimulus alphabet under degraded conditions and to explore the resulting frequencies of confusions and correct responses. Many important psychological concepts may be embedded in mathematical models and tested against such data.

Models centered around feature processing ideas permit sophisticated interstimulus similarity dynamics and have been reasonably successful. A class of models called "descriptive models" make fewer processing assumptions, but one of them, the similarity choice model (see Luce, 1963; Townsend, 1971b, Townsend & Landon, 1982), has performed very well overall. An investigation of descriptive models in the context of the present experiment will appear in a later paper. The reader is referred to Townsend and Landon (1983) for a survey and taxonomy of a wide variety of mathematical models of recognition.

Here we confine our attention to the currently most important feature models. Featural qualities of a stimulus have received the lion's share of attention in the relevant literature. Apart from psychological and physiological experiments, the idea that simple geometric aspects of alphabetic characters play a strategic role in their perception has considerable psychological appeal. Subjectively, they are what we

notice and how linguists often conceive of structures of alphabets (e.g., Watt, 1980).

Earlier studies typically found reasonable fits with feature models (e.g., Geyer & DeWald, 1973; Rumelhart & Siple, 1974). However, the fundamental assumptions of the models were not tested directly and often a given model was not tested against alternative models. Recently, more direct experimental testing of the basic assumptions underlying feature analysis has begun. Certain of these psychologically strategic assumptions have not passed the tests (Townsend & Ashby, 1982; Townsend, Hu, & Ashby, 1981; Wandmacher, 1976). However, aside from the intuitive validity of lines, curves, and angles as features, some of the previously assumed important properties of these have received support. The model parameters that relate to feature perception also behave in a regular way in confusion experiments. Such behavior is not only of interest in its own right, but is also critical in a number of models concerned with higher order processes, such as reading. A model of word perception might fail or succeed simply due to the presence of empirically incorrect assumptions at the featural level.

In the present study, the intent was to find out how a set of simple geometric operationally defined features interact in pattern perception, and to provide further direct tests of feature processing assumptions. The process of extracting simple features from a stimulus will be called the feature sampling process.

The typical notion in a feature processing model is that the features sampled in the sensory process are matched against individual sets of features in memory that make up the set of possible stimuli. A subset of potential response alternatives is formed

(often called the "confusion" or "candidate" set; e.g., Geyer & 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. 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. This set of probabilities is typically assumed to be formed by dividing the response bias of each alternative by the sum of the response biases of all members of the confusion set.

Townsend et al. (1981) have specified a number of typical assumptions presumably underlying the sampling process. We relist them here for the reader's convenience.

Feature Sampling Assumptions

(1) Presence or absence of a feature in a stimulus: A specific feature is either contained in a stimulus in a 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 from a stimulus pattern can be written as the product of the separate probabilities of each feature being sampled.

(5) Across-stimulus invariance: This assumption means that the average (marginal) 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, relative to the range of features in the experiment.

(6) Across-feature invariance: Here it is supposed that all features within a given stimulus possess the same average (marginal) 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 detectability 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."

The first three assumptions are usually taken as implicit in modeling feature processing mechanisms. They probably hold to an acceptable degree in the highly controlled domain of the present and similar studies. The last four are more suspect even in single letter recognition studies. All four have begun to receive direct empirical test over the last few years.

These four assumptions are of psychological import, but sampling independence is of especial concern because, if it fails, not only is the potential number of free parameters greatly magnified, but also the joint sampling probability is no longer a multiplicative combination of independent terms. Sampling independence has been universally assumed in mathematical modeling at least since the early 1970s (e.g., Geyer & DeWald, 1973; Rumelhart, 1971). Garner and Morton (1969) seemed to be referring to this aspect of processing in their "state independence."

As an example, suppose a stimulus, S, is presented and that it is composed of the features a, b, but not c, where a, b, and c are the features out of which all stimulus patterns are formed. Then the probability that, say, feature a is correctly sampled but b is lost while c is incorrectly gained (as a ghost feature) is:

$$P_s(a \ \& \ not \ b \ \& \ c) = P_s(a) \times P_s(b \ lost) \times P_s(c),$$

where $P_s(b \ lost) = 1 - P_s(b)$. Note that, in general, these probabilities can depend on what stimulus was presented (allowing across-stimulus variability and therefore violating Assumption 5). Furthermore, it may be that $P_s(a) \neq P_s(b)$, so that Assumption 6 may be violated. $P_s(c)$ is, by definition, the probability of sampling c as a ghost feature because S does not contain c, and if S' is a stimulus that does contain c, then typically we expect $P_{S'}(c) > P_S(c)$. If $P_S(c) > 0$, the high-threshold assumption is wrong. The important point is that these properties may or may not hold, completely separate from the sampling independence question.

The failure of sampling independence may be expressed either as

$$P_s(a \mid b) \neq P_s(a)$$

or as

$$P_s(a \& b) = P_s(a \mid b) \times P_s(b) \neq P_s(a) \times P_s(b).$$

The probabilities on the far right-hand side of the equation are the marginal or average sampling probabilities.

Such experiments as that on the object superiority effect by Weisstein and Maguire (1978) and the importance of context in letter identification (e.g., Garner & Haun, 1978; Schendel & Shaw, 1976) suggest that the stimulus and response milieu can be influential. However, there have been few studies that attempt to simultaneously address Assumptions 4, 5, 6, and 7 in visual feature confusion experiments and very few that employ mathematical models to vigorously test the assumptions. Most of the previous studies have also been incapable of answering certain questions either because of limitations in experimental design or the analyses.

When the fundamentals of feature processing are studied, certain laboratory conditions may be imposed which are not met in ordinary alphabets. For instance, as Townsend and Ashby (1982) note, it can be valuable to form an artificial alphabet composed of all subsets of the atomic set of features, including the blank stimulus. This permits certain analyses of the sensory and bias processes and minimizes untoward interactions of these processes. Preferably, the models employed should be general enough to allow tests of important assumptions rather than the simple assumption of their truth.

Furthermore, models of higher order processes that depend on feature perception often make some or all of the above highly constraining assumptions (e.g., Rumelhart & Siple, 1974). Thus, we should learn whether and to what extent such assumptions may hold even in the rarified paradigm we address here.

One aspect of feature processing that had received no attention until the Townsend et al. (1981) study was that each feature is in essence a signal and therefore may be associated with a sensitivity, or d', value and a bias or response criterion value. Most previous work has, by fiat, assumed number 7—the high-threshold notion—without testing it and therefore precluding a signal detection analysis. Therefore, a finding that, say, a feature's sampling probability falls as the complexity of the pattern in which it is embedded grows is ambiguous until one learns whether the effect is of sensory or bias nature. We shall refer to signal detection analysis of this type as "macroanalysis."

A critical related aspect that has not been heretofore noticed, is that when dependencies are found in the sampling probabilities they also may be potentially due to a change in bias or sensitivity because of the detection of another feature. Therefore, we designed a data probe technique that evaluates whether sampling dependencies are caused by sensitivity or bias effects. We refer to this as a signal detection microanalysis.

Our present experiment is based on all combinations of three straight lines and one curved segment,

operationally defined as features. The resulting stimulus patterns were designed to be letter-like but unlikely to be confused with real letters.

Finally, in an effort to learn if perceptual set could influence sample dependence, two experimental subgroups were formed. Two observers learned relatively suggestive gestalt names for each stimulus pattern, whereas two other observers were to be trained to report by feature. For instance, the gestalt-report group would be taught to respond "sickle" to a presentation of the pattern ⌐, whereas the feature-report group would report features 2 and 3 (see Figure 2 below). We now turn to a more detailed discussion of the feature models and how they may be implemented.

Feature Process Models

Direct report model. The direct report model assumes that a set of features is simply sampled from the stimulus display and reported as seen. The sampling may be, in principle, independent, partially dependent (some features may be independently sampled, others dependently sampled), or entirely dependent. The magnitude of the sampling probabilities are permitted to depend on the complexity of the stimulus pattern and on the particular feature. Thus across-stimulus and across-feature as well as sampling dependence variability were allowed in the most general model. The high-threshold assumption was also relaxed. Note that comprising the stimulus alphabet of the full set of all possible combinations encourages the use of a direct report strategy because the observer knows that any possible sampled set of features could have been a stimulus. A strong version of the direct report model, which assumed sampling independence, provided the best fit of the Townsend et al. (1981) data.

Biased feature models. The roots of the present feature models lie in the multicomponent model of Rumelhart. However, the general biased feature model is a substantial generalization of the multicomponent model.

The original model follows the overall conception introduced earlier: a set of features is sampled from those contained in the stimulus; each has the same probability of being sampled. After that, a confusion set is formed from all alternatives sufficiently similar to the sample. Then the probability of a certain response from the confusion set is given by the relative response bias strength of that alternative to the sum of all the bias strengths in the confusion set. The original Rumelhart model assumed that features could be lost but not incorrectly sampled (high-threshold assumption), and the confusion sets were not permitted to include alternatives that lacked features contained in the sample.

Furthermore, the bias probabilities were presumed to follow Bayesian principles in the first version of the model. All but one of these constraints have been

dropped in the present models. Because of problems in convergence for the complex computer routines, it was unfortunately necessary to retain the high-threshold assumption in the present biased feature model. Observe that if the only stimulus alternative that is permitted to be in the confusion set is that corresponding to the sampled set of features, the high-threshold direct report model is derived as a special case.

It should be pointed out that the macro- and micro-analyses mentioned above are based on the direct report model.

METHOD

Observers

Four observers took part in the experiment. Three of them did so as part of an upper division experimental psychology course; the fourth (Observer 1) was the second author. Three of the observers possessed normal vision. Observer 1 possessed an astigmatism that degraded his perception of the horizontal line. (The astigmatism was undetected at the time of the experiment.)

Apparatus

A Gerbrands two-field tachistoscope (Model T-2b) was used to present the stimuli. A stimulus set was formed by using all combinations of three equal-length lines (33.4 min visual angle) and one semicircle segment (33.4 min visual angle for diameter) as features. These segments are denoted as: left vertical line (designated 1), open vertical curve (designated 2), top horizontal line (designated 3), and major diagonal line (designated 4), as shown in Figure 1. Thus, the stimulus set consisted of $2^4 - 1$ nonblank stimuli plus one blank stimulus described by the absence of all four segments, as shown in Figure 2. Each feature was presented the same number of times overall, inasmuch as each feature is equally represented in the stimulus set.

A prestimulus fixation field was represented by a set of four dots that 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 side subtended an angle of about 2 deg at the observer's eyes. The stimulus was drawn in dark ink on a white card. The luminance of the prestimulus field was 35.63 cd/m², and the luminance of the stimulus field was maintained at about 33.06 cd/m². The stimulus durations were 8, 6, 6, and 7 msec for observers 1, 2, 3, and 4, respectively. Two types of responses were given. The first pair of observers were instructed to respond in numbers (e.g., nothing, 1, 2, 3, 4, 12, 13, 14, 24, 23, 34, 123, 124, 234, 134, and 1234). The other two observers learned and re-
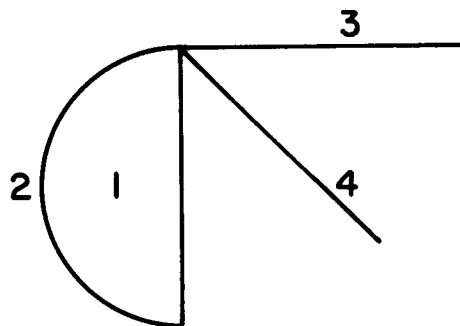
STIMULUS   RESPONSE

STIMULUS	Number	Name
blank (∅)	nothing	nothing
I	I	vertical
C	2	curve
—	3	horizontal
\	4	oblique
◁	12	bow
⌐	13	book
⋀	14	top
⌒	24	fishhook
⌜	23	sickle
⋎	34	punch
⌐	123	club
⋀	124	handle
⋎	234	extinguisher
⋀	134	envelope
⋀	1234	map

Figure 2. Stimulus-response paradigm.

sponded with names (see Figure 2). Responses were given verbally and were recorded by the experimenter on a recording sheet.

Procedure

All observers had an hour of practice for 5 days, which yielded a combined total of 5 h practice for each observer. Before the experiment began, the stimulus duration was set for each observer individually, so the probability of being correct was approximately 50%. Twenty trials of practice preceded each of 15 experimental sessions for each observer. Thus, over 15 experimental days, each observer was presented each stimulus 300 times. Thus, sufficient data were gathered to permit relatively powerful statistical conclusions to be drawn on an individual observer or group level.

RESULTS

The 16 × 16 confusion matrices obtained from each observer were summed over the 15 experimental sessions and are given in Tables 1-4. The analyses are divided into two parts: (1) The first section includes statistical fits of various direct report feature models (which predict varying degrees of dependencies between the features), analysis of report-instruction differences (i.e., Observers 1 and 2 reported features of a stimulus, whereas Observers 3 and 4 reported gestalt-like names for stimuli), and the biased feature models, which include decision parameters. (2) The second section, which probes more deeply into the specific nature of the structure within the data, includes discussion of ghost feature sampling, feature dependencies, analysis of feature sensitivity, and bias in report, through signal detectability theory.

Figure 1. Stimulus configuration.

Table 1
P(R_j | S_i), the Estimated Probability of Reporting R_j Given Stimulus S_i in a Confusion Matrix for Observer 1

R_j

$\hat{P}(R_j\|S_i)$	φ	I	(	−	\	◁	Γ	Ν	∧	Ϲ	≺	ᒋ	∧	Ϲ	Ϝ	Ϛ
φ	.917	.013	.010	.027	.020	0	0	.003	0	.003	.007	0	0	0	0	0
I	.027	.747	.010	.003	.003	.003	.097	.073	.007	.003	.003	0	.003	0	.020	0
(	.057	.023	.693	.003	.013	.013	.013	.013	.053	.077	0	.017	.007	.017	0	0
−	.620	.010	.007	.277	.017	.003	.013	.007	0	.007	.030	.003	0	0	.007	0
\	.183	.007	.013	.050	.667	0	.003	.003	.007	.003	.053	.003	0	0	.003	.003
◁	.020	.050	.073	0	.003	.643	.013	0	.010	.013	0	.103	.043	.007	.007	.013
Γ	.013	.293	.020	.020	.003	.003	.550	.047	0	.003	.003	.017	0	.007	.020	0
Ν	.020	.177	.013	.003	.043	.017	.070	.500	.013	0	.013	0	.003	.010	.117	0
∧	.047	.007	.150	.010	.050	.010	.013	0	.400	.083	.003	.007	.017	.147	.003	.003
Ϲ	.033	.007	.367	.030	.003	.010	.023	.003	.023	.453	0	.017	.010	.013	.003	.003
≺	.130	.013	.020	.097	.360	0	0	.013	.007	.013	.337	0	0	.003	.007	0
ᒋ	.007	.023	.057	.013	0	.320	.060	.017	.013	.050	.003	.330	.030	.013	.013	.050
∧	.020	.017	.050	.003	.017	.227	.013	.023	.050	.013	0	.100	.350	.047	.003	.067
Ϲ	.033	.003	.097	.020	.050	.010	.010	.003	.237	.157	.023	.027	.003	.307	0	.020
Ϝ	.013	.057	.020	.007	.020	.007	.210	.250	.027	.003	.017	.003	.003	.003	.357	.003
Ϛ	.003	.003	.020	.007	.003	.143	.023	.010	.017	.040	.007	.237	.183	.043	.023	.237

S_i

Table 2
P(R_j | S_i), the Estimated Probability of Reporting R_j Given Stimulus S_i in a Confusion Matrix for Observer 2

R_j

$\hat{P}(R_j\|S_i)$	φ	I	(	−	\	◁	Γ	Ν	∧	Ϲ	≺	ᒋ	∧	Ϲ	Ϝ	Ϛ
φ	.827	.020	.023	.033	.057	.007	.003	.007	.013	.007	.003	0	0	0	0	0
I	.120	.630	.017	.023	.037	.023	.043	.053	.010	.010	.007	.010	0	.003	.010	.003
(	.143	.057	.467	.030	.047	.040	.020	.003	.053	.087	0	.017	.010	.023	.003	0
−	.197	.003	.007	.607	.023	0	.020	0	.003	.020	.090	.010	0	.007	.010	.003
\	.103	.013	.010	.027	.730	0	0	.017	.020	.003	.063	0	.013	0	0	0
◁	.110	.067	.120	.007	.040	.417	.017	.017	.037	.030	.003	.073	.050	.003	0	.010
Γ	.040	.070	.007	.137	0	.003	.530	.023	0	.050	.030	.047	.003	.010	.047	.003
Ν	.033	.143	.020	.017	.187	.010	.023	.397	.053	0	.010	.003	.027	.017	.057	.003
∧	.037	.017	.153	.010	.137	.023	.003	.027	.440	.027	.023	.003	.027	.063	0	.010
Ϲ	.037	.003	.057	.107	.010	.020	.043	0	.003	.553	.013	.093	.003	.053	0	.003
≺	.050	.003	.013	.127	.097	0	.013	0	.007	.017	.587	.013	0	.007	.057	.010
ᒋ	.013	.030	.030	.070	0	.067	.090	.013	.007	.180	.027	.397	.003	.033	.017	.023
∧	.040	.017	.070	.013	.070	.120	.020	.050	.170	.017	.007	.070	.237	.033	.010	.057
Ϲ	.027	.007	.027	.037	.037	0	.020	.007	.073	.167	.110	.040	.003	.400	.013	.035
Ϝ	.017	.017	.010	.023	.020	.010	.143	.070	.020	.030	.123	.010	.010	.037	.443	.017
Ϛ	.010	0	.027	.043	.017	.023	.027	.007	.027	.067	.067	.227	.053	.137	.060	.210

S_i

Table 3
P(R$_j$ | S$_i$), the Estimated Probability of Reporting R$_j$ Given Stimulus S$_i$ in a Confusion Matrix for Observer 3

R$_j$

$\hat{P}(R_j \mid S_i)$

S$_i$	φ	I	(	−	\	⊄	Γ	⋀	⋀	⌐	⋜	⋤	⋀	⋉	⋈	⋈
φ	.797	.023	.033	.033	.047	0	.003	.013	.010	.010	.023	0	0	.003	0	.003
I	.233	.583	.027	.007	.013	.023	.053	.040	.003	0	.003	0	.010	0	.003	0
(	.190	.033	.510	.003	.033	.080	.017	.020	.017	.030	.013	.013	.020	.007	.013	0
−	.283	.0	.020	.433	.010	.010	.030	.007	0	.033	.133	.020	0	.007	.010	.003
\	.170	.007	.013	.023	.680	.017	0	.017	.023	.003	.033	0	.007	.003	0	.003
⊄	.103	.067	.053	.010	.017	.537	0	.027	.023	.007	.010	.040	.077	.007	.007	.017
Γ	.087	.093	.013	.127	.013	.023	.467	.017	.003	.027	.040	.033	.003	.003	.050	0
⋀	.107	.107	.023	.013	.113	.017	.010	.477	.017	.003	.027	.003	.023	.003	.047	.010
⋀	.097	.007	.087	.007	.073	.037	.013	.043	.420	.023	.017	.023	.060	.077	.003	.013
⌐	.067	.003	.087	.090	.020	.020	.060	0	.007	.480	.023	.077	.010	.023	.020	.013
⋜	.100	.013	0	.110	.177	0	.017	.010	.003	.003	.483	.010	.007	.013	.043	.010
⋤	.067	.020	.017	.030	.017	.100	.053	.013	.017	.117	.013	.383	.047	.023	.010	.073
⋀	.083	.017	.017	0	.073	.080	.007	.030	.117	.020	.020	.027	.397	.033	.010	.070
⋉	.033	.007	.043	.030	.027	.007	.040	.003	.117	.127	.093	.033	.027	.387	.010	.017
⋈	.030	.030	.023	.007	.060	.007	.110	.090	.013	.023	.090	.020	.010	.020	.437	.030
⋈	.033	.013	0	.017	.007	.070	.023	.013	.043	.023	.047	.127	.073	.093	.043	.373

Table 4
P(R$_j$ | S$_i$), the Estimated Probability of Reporting R$_j$ Given Stimulus S$_i$ in a Confusion Matrix for Observer 4

R$_j$

$\hat{P}(R_j \mid S_i)$

S$_i$	φ	I	(	−	\	⊄	Γ	⋀	⋀	⌐	⋜	⋤	⋀	⋉	⋈	⋈
φ	.867	.027	.003	.040	.040	0	0	.013	0	.003	.007	0	0	0	0	0
I	.137	.737	.010	.007	.013	.003	.037	.040	.007	0	.003	0	0	0	.007	0
(	.333	.073	.437	.017	.020	.037	.017	.007	.027	.010	.010	.007	0	.003	.003	0
−	.337	.013	.010	.513	.017	.007	.030	.003	0	.013	.050	.007	0	0	0	0
\	.153	.027	.010	.010	.733	.003	0	.017	.010	0	.033	0	.003	0	0	0
⊄	.167	.183	.090	.007	.010	.423	.020	.037	.003	.007	.010	.017	.027	0	0	0
Γ	.103	.127	.010	.090	.003	0	.577	.013	0	.017	.013	.007	.003	.003	.033	0
⋀	.113	.130	.017	.007	.150	.007	.013	.507	.017	0	.007	.003	.003	0	.027	0
⋀	.153	.013	.107	.010	.120	.010	.007	.067	.420	.007	.013	.007	.033	.017	.013	.003
⌐	.120	.007	.127	.080	.007	.013	.083	.003	.003	.453	.020	.047	0	.030	.007	0
⋜	.137	.013	.017	.120	.160	.003	.020	.013	.007	0	.487	.003	.003	.003	.013	0
⋤	.097	.070	.017	.060	.013	.100	.163	.010	0	.053	.023	.353	.003	.007	.010	.020
⋀	.120	.053	.030	.013	.070	.130	.010	.117	.093	0	.017	.017	.290	.003	.013	.023
⋉	.073	.003	.033	.067	.057	0	.033	.013	.063	.143	.097	.017	.003	.350	.027	.020
⋈	.033	.060	.010	.013	.040	.007	.163	.110	.017	.007	.100	.013	.003	.003	.420	0
⋈	.060	.027	.037	.033	.017	.043	.067	.050	.020	.060	.067	.117	.080	.073	.107	.143

Before further discussion, a final note should be made concerning the performance of Observer 1. As can be seen from the data in Table 1, Observer 1's astigmatism affected his perception of the horizontal line or feature. However, this handicap caused very little perturbation in the analyses and conclusions; the details of his perceptual behavior turned out to be very close to those of the other observers in most instances.

Analysis of Feature Model Predictions

Direct report model fits. We now investigate the sampling independence question using the direct report feature model as our structure for inquiry. We employ a shorthand notation for simplicity's sake. For instance, (1, 2, 3, 4) represents the complete independence of elements 1, 2, 3, and 4. (12, 3, 4) denotes a decomposition of the stimulus configuration into three independent elements, 12, 3, and 4, and permits a partial or complete statistical dependency to exist between the two features 1 and 2, but not between 3 and 4 or between the pair 12 with 3 or 4. This notation is used throughout the rest of this section. The dependence denoted by any one model can in principle be in the form of a positive or negative correlation as reflected in the conditional sampling probabilities of the two (or more) features. Note that a complete positive correlation between a pair of features would mean that both are sampled or rejected simultaneously, whereas a complete negative correlation would imply that if one is sampled, on a given trial, the other is rejected. A more intricate and revealing signal detection analysis will be developed below.

In all, a total of 14 direct report feature models were tested. These models included total independence of features (1, 2, 3, 4), all combinations of three independent feature structures, (12, 3, 4), (13, 2, 4), (14, 2, 3), (23, 1, 4), (24, 1, 3), (34, 1, 2), all combinations of two independent feature structures, (12, 34), (13, 24), (14, 23), and all combinations of two independent feature structures with asymmetric contributions (123, 4), (124, 3), (134, 2), (234, 1).

All these models were tested under the most liberal assumptions, that is, across-stimulus and across-feature sampling variability and low-threshold feature sampling (ghost-feature sampling capability), so that, for instance, sampling independence would not be rejected for the wrong reason.

In the case of the direct report model with complete independence (i.e., the (1,2,3,4) model), estimation and fit were accomplished through both a maximum likelihood procedure and a minimum chi-square procedure. Both methods falsified this model at the $\alpha = .01$ level for all observers. Of all the models that permitted featural dependencies, the best were those with a positive correlation between the curve and the

vertical line. However, even these models had to be rejected statistically at the $\alpha = .01$ level.

We are forced to conclude that within a direct report framework, dependencies exist throughout all features present in a pattern. What type of dependencies are there? Table 5 shows the joint probabilities versus the product of the marginal probabilities for various feature clusters contained in a stimulus pattern. Table 6 performs the same service for ghost feature combinations. Positive dependencies (correlations) are indicated by an observer's left column (the joint probability) being larger than the right column (the product of the marginal probabilities). Negative dependence is revealed by the reverse inequality, and sampling independence is revealed by equality.

Out of the 80 combined comparison pairs of Tables 5 and 6, there are only 9 that suggest a negative dependence. A sign test yielded significance of the positive dependencies at the $\alpha = .05$ level. Furthermore, only Observer 1 failed to reach significance in a paired t test at $\alpha = .05$. The three normally sighted individuals clearly exhibit a reliable positive sampling dependence. Thus, if it is known that one feature has been sampled, whether it be real or ghost, the likelihood that another feature will also be sampled is increased. A question that has not been previously addressed in the literature is whether such a correlation proceeds from alterations in sensory sensitivity or feature-response bias. This issue will be taken up below in the micro-analysis section.

Table 5
The Average Joint Probabilities of Sampling Two or Three Features Versus the Product of the Average (Marginal) Probabilities of Sampling Features Contained in the Stimuli ("Real" Features)

Joint Probability	Product of Marginal Probability	Feature Group O_1		Feature Group O_2		Gestalt Group O_3		Gestalt Group O_4	
P(Ɑ)	vs P(│)P(C)	.769	.758	.509	.468	.622	.551	.447	.388
P(Γ)	vs P(│)P(‾)	.533	.523	.572	.564	.558	.524	.548	.512
P(Ͷ)	vs P(│)P(\ )	.532	.530	.468	.437	.522	.509	.473	.448
P(C)	vs P(C)P(‾)	.499	.499	.654	.639	.592	.567	.472	.431
P(⋀)	vs P(C)P(\ )	.532	.520	.493	.495	.579	.541	.409	.374
P(⋁)	vs P(‾)P(\ )	.347	.357	.578	.570	.548	.522	.478	.447
P(Ɑ̄)	vs P(│)P(C)P(‾)	.427	.438	.428	.394	.478	.412	.317	.261
P(⍒)	vs P(│)P(C)P(\ )	.418	.408	.265	.277	.457	.387	.268	.221
P(Ʀ)	vs P(│)P(‾)P(\ )	.310	.321	.365	.366	.442	.393	.335	.311
P(Ϲ⋁)	vs P(C)P(‾)P(\ )	.303	.304	.390	.392	.435	.397	.293	.256
Ave.		.467	.466	.472	.460	.523	.480	.404	.395

Note: Each probability is averaged across stimuli that contain each pair or triple of features. (O = observer)

Table 6
The Average Joint Probabilities of Sampling Features Versus the Product of Average Marginal Probabilities of Sampling Features That Are Not Contained in the Stimuli ("Ghost" Features)

Joint Probability	Product of Marginal Probability	Feature Group O_1		O_2		Gestalt Group O_3		O_4	
P(⊐) vs	P(∣)P(C)	.003	.001	.014	.003	.022	.005	.008	.002
P(⌐) vs	P(∣)P(⁻)	.018	.008	.015	.012	.027	.015	.014	.006
P(Ν) vs	P(∣)P(\ )	.014	.004	.011	.011	.033	.016	.009	.007
P(⊓) vs	P(C)P(⁻)	.007	.005	.015	.008	.012	.006	.002	.001
P(∧) vs	P(C)P(\ )	.004	.002	.015	.009	.012	.010	.003	.002
P(⊼) vs	P(⁻)P(\ )	.018	.010	.018	.016	.028	.011	.011	.004
P(⊓) vs	P(∣)P(C)P(⁻)	.001	–	.0	–	.003	–	.0	–
P(∧) vs	P(∣)P(C)P(\ )	.0	–	.002	–	.003	.001	.0	–
P(Ν) vs	P(∣)P(⁻)P(\ )	.0	.001	.002	.002	.008	.001	.002	–
P(⊼) vs	P(C)P(⁻)P(\ )	.0	–	.003	.001	.003	–	.0	–
Ave.		.006	.003	.010	.006	.015	.007	.005	.002

Note: – *denotes the value greater than 0 and less than .001. Each probability is averaged across stimuli that do not contain the pair or triple of features.*

Gestalt versus feature reports. Recall now that Observers 3 and 4 were taught to respond to the stimuli by reporting names for each stimulus (gestalt-report group), whereas Observers 1 and 2 were taught to respond by naming the features they saw on each trial (feature-report group).

The likelihood ratio statistics computed from the direct report models showed that for every model the numerical value of the fit was better for the feature-report group (i.e., Observers 1 and 2). Moreover, Tables 5 and 6 reveal an even stronger positive dependence for the gestalt-report group than for the feature-report group. These differences are statistically significant ($\alpha = .01$), as manifested in sign tests that were carried out. Both results suggest that features tended to be sampled or not sampled together more by the gestalt group more than by the feature group.

The above analyses and the fact that none of the direct report models provided statistically adequate fits for either group tell us that although the report instructions succeeded in affecting sampling correlations, strong featural dependencies remain even in the feature-report group. A trend analysis across sessions showed that the feature- vs. gestalt-report-group differences were present early in the experiment and remained unchanged over days.

Biased feature models. The above evidence disconfirms direct report models with partial dependencies or independence as complete accounts of the recognition process in this experiment. However, it is possible that an intervening decision process perturbs the reports in such a way as to obscure true underlying feature sampling independence. All the models in the present group assume sampling independence at the sensory level followed by a biased decision process.

First, the multicomponent model (Rumelhart, 1971) with a generalized non-Bayesian decision process was falsified [$\chi^2(223) > 3,000$ for each observer]. Next, more general models were tested. These allowed across-stimulus and across-feature variability in the sampling parameters. However, the high-threshold assumption was maintained. Such models provided the best feature process explanations in an earlier study, but they were not the best overall models (Townsend & Ashby, 1982) and are closely related to those employed in several cases by Wandmacher and his associates (e.g., Wandmacher, 1976; Wandmacher, Kammerer, & Glowalla, 1980). These also failed in the present data, although they were not as inadequate as the multicomponent model [$\chi^2(192) \geqslant 800$].

The hypothesis of an independent feature sampling phase followed by a response bias phase must be considered very unlikely, in view of the above findings. It is, of course, possible that a partial dependence feature phase that permits acquisition of lost features, is followed by a response bias phase. However, without an independent specification of the feature dependencies, such a model is too cumbersome and unparsimonious to be interesting, and in fact leads to nonconverging parameter estimation interactions in the present data.

Detailed Analyses of Feature Structure

It is assumed throughout the following presentation that a feature is reported if, and only if, it is sampled; this is the direct report model. However, we term the sampling of a feature to be a process that depends both on sensitivity and feature-strength aspects as well as decision and bias factors. This portion of the paper will then address two major topics, both in the context of the feature reports for the four observers:

(1) Across stimulus context effects on sampling probabilities, as well as d′ and β in feature processing at the overall, or *unconditional*, level—called the "macroanalysis."

(2) A fine-grain analysis of feature sampling dependencies also employing signal detectability theory, but this time on a *conditional* sampling basis—called the "microanalysis."

Roughly, an unconditional feature d′ in (1) represents the feature sensory sensitivity for a given feature when it is embedded in a given stimulus pattern. In contrast, a conditional d′ is calculated given that, say, another feature has been correctly sampled. Although logically separate, a specific system might interlink the unconditional and conditional statistics. As an example, suppose that a feature-processing system evidenced lateral interference among features

in a stimulus pattern, and that this was manifested in a decreased d′ for any particular feature when another feature was correctly sampled. Then that type of lateral interference could (but does not have to) grow more severe, the more features that a stimulus contained. In such a system, the two types of d′ would be intimately related through the system-processing mechanisms. This example provides a handy comparison for our actual results below. The macroanalysis develops the unconditional statistics, and the microanalysis develops the conditional statistics.

Signal detection analysis: Macroanalysis. One of the striking results in Tables 1-4 is that the probability correct (diagonal entries) decreases consistently with the increase of the number of features in the stimulus. Now, if the hypothetical feature processor checked all feature positions with equal capacity devoted to presence versus absence of a feature, then, other things being equal, performance should not suffer as the number of features in the stimulus increased. However, earlier results indicate that other things are not equal; that, in fact, features are lost more readily than ghost features are added (see Garner & Haun, 1978; Townsend & Ashby, 1982; Townsend, Hu, & Ashby, 1980). In this case, given that the probability of falsely sampling ghost features is lower than the probability of losing existing features in the stimulus, the stimuli with fewer features should show less deterioration on the average in the percept. A stimulus with many features would have a larger probability of being confused with the alternatives containing subsets of the features contained in the original stimulus. This mechanism predicts the qualitative order of decrement in the diagonal probabilities viewed in Tables 1-4.

On the other hand, in order to learn if the predicted decrements are sufficient to explain the total degradation, it is necessary to probe the individual real and ghost feature sampling probabilities. The fact that more features are lost when a pattern is exposed than there are ghost features that are falsely gained corresponds to an ordinary signal detection situation in which the frequency of missed signals is greater than the frequency of false alarms. Under the standard assumptions of signal detectability theory, this implies that the decision criterion is to the right of the intersection of the signal + noise curve and the noise-alone curve. Then, even if the criterion and the sensitivity are unchanged when patterns with more features are presented, one would still expect a greater loss of features. It is therefore necessary to compute d′ and β over alteration in number of features in the stimuli in order to ascertain whether a true loss of sensitivity is occurring. If the d′s on individual features decline, then we have evidence of true sensory degradation.

Table 7 shows the group average sampling probabilities and standard errors for the four different features, given the possible stimulus representations. The correct sampling probability for any feature in Table 7 is defined as the probability of the specific feature being included in the response, given the feature was contained in the stimulus on a particular trial. The general trend is that for the inside features (features 1 and 4), the sampling probabilities decrease as the number of features in the stimulus increases. The outside features (features 2 and 3) show the opposite behavior in that their correct sampling probabilities increase with an increase in the number of features in the stimulus. Trend analyses reveal significant linear trends ($\alpha = .01$) for each feature, decreasing for the inside features and increasing for outside features.

The changes across stimuli in sampling probabilities could be due, in principle, either to a shift in decision bias or an alteration in sensitivity, or both.

Table 8 reports the d′ and β statistics; the appendix relates the details of these computations. The general impressive decrease in d′ with an increase in the number of features in the stimulus indicates a decrease in feature perceptibility (sensitivity) as the number of component features increased. On the other hand, we now see that the increase in the real feature sampling probabilities of the outside features 2 and 3 (see Table 7) with the increase in number of features in the stimulus is not explained by a greater sensitivity to these features, but rather by a dramatic decrease in β. The decline of β seems to denote a greater ten-

Table 7
Group Average and Standard Error of Feature Sampling Probabilities Estimated by the Maximum Likelihood Estimate Method for Four Different Features Contained in Stimuli With Different Numbers of Features

| | | Feature | | | |
		I	⊂	—	\
one-feature	M	.813	.693	.597	.786
stimuli	SE	.022	.027	.028	.024
two-feature	M	.753	.750	.708	.713
stimuli	SE	.014	.014	.015	.015
three-feature	M	.725	.745	.708	.672
stimuli	SE	.015	.014	.015	.016
four-feature	M	.709	.766	.717	.588
stimuli	SE	.026	.024	.026	.028

Note.*One-feature stimuli are I , ⊂ , — , and \ .
Two-feature stimuli are ◖ , ⌐ , ∧ , ◗ , ⊏ , and ⌐ .
Three-feature stimuli are ⊓⁻, ◖∧, ◗⊏, and ⌐ .
Four-feature stimulus is ◖⊏ .

Table 8
Group Average Unconditional d' and β Over Four
Observers for Each Feature

	Feature							
	⏐		⊂		—		\	
	d'	β	d'	β	d'	β	d'	β
one-feature stimuli	2.80	3.30	2.52	8.06	1.90	3.46	2.33	2.62
two-feature stimuli	2.25	2.63	2.41	4.91	2.00	1.95	1.89	2.21
three-feature stimuli	1.88	1.86	2.21	2.62	1.77	1.84	1.78	2.30
four-feature stimuli	1.79	1.70	2.08	1.75	1.50	1.34	1.32	1.86

dency to give positive feature responses in the presence of weaker sensory states.

The fact that the d' statistic for a single feature decreases with an increase in the number of features contained in the stimulus indicates that the feature sampling is limited-capacity (in an operational sense) in this experiment. Interestingly, the decrease in d' is still less for the outside features, (and —, than for the inside features, ⏐ and \. These effects will be discussed further below. We now turn to a closer look at the sampling dependency phenomena.

Signal detection analysis: Microanalysis. An important question, when considering the mutual effects of features on the perception of one another, is the effect of the actual sampling of a particular feature on the sensitivity and bias towards reporting another feature. We learned earlier from positive dependencies in the feature sampling probabilities that the (within-trial) probability of reporting, say, feature "i" went up if another feature, "j," was known to be sampled (correctly reported to be present). However, we do not yet know whether that was due to sensory or bias effects. Note that the macroanalyses of the section above cannot answer this question because they yield information only about the average overall effects that features have on one another; they cannot reveal anything about what happens on a within-trial basis.

The parameters d' and β must be computed for each feature conditioned on (1) the sampling of another feature in a particular stimulus versus (2) the failure to sample another feature, when both features are assumed to be contained in the stimulus in either case. Recall that we are basing this discussion on the precept that sampling implies report, and vice versa. However, the analyses could be viewed in an operational sense as conditioning on "report," rather than "sampling." We wish to compute the following statistics for all possible combinations of features i and j, i ≠ j, and all possible stimuli k, where features i and j are contained in stimulus k.

d'(feature i | stimulus k was presented and feature j was sampled)

β(feature i | stimulus k was presented and feature j was sampled)

d'(feature i | stimulus k was presented and feature j was not sampled)

β(feature i | stimulus k was presented and feature j was not sampled)

The process of computing the d' and β described above revolves around obtaining the probability of a hit and the probability of a false alarm for each of the features under the conditions specified by the particular d' and β. Once these values are computed, the ordinary signal detection methods are applied. Under the conditions supplied above, we computed the possible 48 conditional d's and 48 conditional βs for each observer.

For economy of space, Table 9 reports only the average results taken across features as well as observers and includes the average unconditional d' and β from the earlier macroanalyses for comparison purposes. The d' statistic in the case of correctly sampling another feature is greater than the average unconditional d' that, in turn, is greater than that conditioned on the failure to sample another feature in the stimulus. Moreover, β also decreases in the same fashion. The trends indicated in Table 9 are highly significant and were true for all individual features. For example, a t test between the unconditional versus positive conditional d' yielded t(27) = 6.05, p < .01. Thus, it may be concluded that featural sensitivity is dramatically elevated when another feature is captured and is diminished when a separate feature in the stimulus pattern fails to be detected. In contrast, correct sampling of another feature is associated with an elevated response criterion β relative to the average or negatively conditioned β. This means that the observer is more prone to report a weak-feature signal in the absence of another feature detection.

Table 9
Values of Conditional d' and β Averages Across
Observers, Features, and Stimuli

	Case					
	1		2		3	
	d'	β	d'	β	d'	β
M	2.18	2.78	1.97	2.41	1.46	2.03

Note—In all three cases, both features are contained in the stimulus. (1) d' and β for each feature is conditioned on the sampling of a certain different feature (positive conditional case). (2) d' and β for each feature is unconditional (marginal case). (3) d' and β for each feature is conditioned on the failure to sample a certain different feature (negative conditional case).

DISCUSSION

Comparison with other studies in the present domain is important but must be accomplished with caution. It is difficult to compare previous studies, even in the studies closest in design to our own. First, other very strong and possibly wrong assumptions were sometimes made which might distort the outcome. Second, some of the conclusions were drawn in the context of a biased feature model, which thereby included a response bias phase, whereas others were tested in the presence of the direct report model (indeed, in some cases the latter model provided the better fit). Third, outside of our laboratory, there has apparently been no application of signal detectability theory to ascertain d's and βs of individual features.

Nevertheless, some general conclusions seem to be emerging from a more or less homogeneous set of experiments across different laboratories. In particular, those of the Wandmacher group at Darmstadt are closest in methodology and experimental design to our own.

Sampling Independence

Sampling independence refers to the probabilities of acquisition of feature subsets, and in a direct report model they are also the feature report probabilities. Wandmacher's laboratory has generally deemed this assumption to be acceptable (called "stochastic independence" in Wandmacher, 1976, and "state independence" in Wandmacher et al., 1980). Their tests were carried out in the context of a model postulating high thresholds and a postsensory decision phase (i.e., a biased feature model).

In one study, we found that sampling independence, even in the very strong direct report model, provided a very good fit to the data (Townsend et al., 1981). That study used two orthogonal lines as features. In contrast, Townsend and Ashby (1982) used more complex letter stimuli with high similarity, but without exposing all combinations of the constituent features, and found widespread violations of independence.

The present results again reject sampling independence, but go beyond previous studies in carrying out signal detection analyses on the conditional sampling probabilities. We termed these "microanalyses," in contrast to the averaged overall d' and β used in testing the other assumptions.

Townsend and Ashby (1982) found evidence of both positive and negative interactions among features on sampling probabilities. In the present work, which, unlike the earlier study, included all combinations of features and the blank stimulus, positive interactions strongly predominated (Tables 5 and 6). The microanalyses revealed that sensitivity actually was strongly affected in a positive manner by the capture of another feature contained in the stimulus.

The decision criterion β was also influenced, showing a marked decline from when another feature had been correctly detected to when another feature was missed from the stimulus.

Overall, it appears that sampling independence can be an acceptable hypothesis in sufficiently rarified circumstances—that is, simple homogeneous alphabets with about the same number of features and the like. It appears that when a large variation in number of features present in a stimulus can occur, the assumption must be discarded. Or, in the typical case in which the stimulus set does not include all combinations of the feature set, decisional influences may enter that modify the set of features actually sampled as a response is selected (Townsend & Ashby, 1982).

Across-Stimulus Invariance

Across-stimulus invariance is soundly rejected by the present data (see, e.g., columns in Tables 7 and 8). The sampling probabilities of the individual features vary across stimuli, but more importantly, the unconditional d' and β evidenced an orderly decrease as the number of features in a stimulus increases. This indicates some type of sensory limited capacity and an alteration in decision criterion as a function of stimulus complexity. Obviously, each stimulus is equally complex in terms of its overall information value and in the information provided by presence or absence of each component feature. Therefore, the actual visual presence of other features must be playing a strategic role.

Wandmacher (1976) found that feature sampling probabilities varied across stimuli even though the basic feature set consisted of only three line-segment features and the actual stimulus alphabet was composed of one or two segments. (He called this assumption "context independence".) The Wandmacher group later made a similar determination with more complex stimuli (Wandmacher et al., 1980). Townsend et al. (1981) employed only two straight lines, either alone or affixed at a right angle, plus the blank stimulus, surely one of the simplest feature alphabets possible. There were two conditions, a connected condition, in which the two lines were physically connected when they were both present, and a gap condition, in which the two lines were separated when they were both present. In the connected condition, d' was invariant across stimuli; in the gap condition, it was not.

Finally, Townsend and Ashby (1982) discovered success of this assumption, not at the sampling probability level but at the d' and β level. That is, β revealed some variation across stimuli but d' was quite invariant. The stimuli they employed were letters composed of three or four straight lines of equal length.

Thus, although feature detectability may be roughly equal across different characters with sufficiently

similar patterns (as in Townsend & Ashby, 1982), violations of this assumption are easily found. It appears that the most likely way to violate across-stimulus invariance is to use a stimulus alphabet with a wide variation in number of features or with a variety of features (e.g., curves vs. angles, etc.).

Some type of lateral interference may be responsible for the decrement in d' as the stimulus complexity rises. The microanalyses of sample dependence argue against the view that the correct detection of a feature causes greater interference for other features on a within-trial basis. For instance, one hypothesis that has been suggested for certain multisymbol paradigms is that a feature channel may be occupied by a similar feature, thus preventing detection of the "intended" feature (Estes, 1972). Our microanalysis results are not supportive of this idea in single-symbol experiments. However, other types of lateral interference may be active.

Across-Feature Invariance

The feature processing parameters also vary from feature to feature even within a single pattern, as indicated in the rows of Tables 7 and 8. For instance, the unconditional d's for (and ⁻ are higher than for | and \. Perhaps this is due to the fact that the first two are more to the left than the others and therefore more compatible with reading habits. The response criteria also differ, although not in a clearly systematic way.

Both of the Wandmacher studies mentioned above (1976; Wandmacher et al., 1980) falsified across-feature invariance. The perceptibility of the horizontal versus the vertical line also differed in the Townsend et al. (1981) experiment. The Townsend and Ashby (1982) experiment found some small variation in d' across features, but most of the variation in sampling probabilities across features was due to large changes in β, and the latter reflected payoffs and the frequency with which a feature appeared in the stimulus letters.

So, again, there are circumstances in which a rough approximation to across-feature invariance may be supported, but in general, even in the context of a quite simple featural alphabet, it may be violated.

High-Threshold Assumption:
Do Ghost Features Exist?

The Wandmacher group has never tested this assumption, and all other feature models that are sufficiently well specified to predict numerical confusion matrices have also simply assumed that features never appear as false alarms.

We have now completed four studies that cast doubt on this assumption. The frequency of false alarms in Townsend et al. (1981) averaged .11. The average relative frequency in the Townsend and Ashby (1982) investigation was .30. The present work dis-

covered many false alarms; note that even the blank stimulus drew feature-present reports about 17% of the time.

Finally, a new experiment (Hu & Townsend, 1984), which varied both payoffs and stimulus energy with an alphabet very similar to the present one, again turned up a high incidence of false alarms in all conditions.

Thus, the high-threshold assumption no longer appears tenable; its inclusion in any feature-processing model should be considered highly suspect.

Conclusions

It may be inferred from the above evidence that the assumptions (sampling independence, across-stimulus and across-feature invariance, and high-threshold) tested in the present and similar studies, to the extent that they hold at all, are tenuous and highly dependent on a highly controlled experimental environment for their existence. In a sense, the surprising thing is that they should hold under any circumstance. The fact that they do may be of real importance in ultimately pinning down just where the higher order complexities arise. Of course, much work remains to be done to determine more precisely the laws of operation of feature mechanisms, or whatever mechanisms give rise to the regular behavior of this and other experiments.

It would be relatively easy to formulate a mathematical model that would do a fair job of fitting the present data, but without a better handle on the aforementioned perceptual mechanisms, it would be a rather indifferent exercise. However, it proves worthwhile to consider some of the possibilities that are capable of producing our results.

A qualitative framework that is consistent with our findings is the following. Suppose a roughly delineated form is available during early stages of processing and this, at least on some of the trials, gives way to a more highly defined subset of features. We may assume that the integrity of this "blob" form has a distribution with a mean that is higher for stimuli with fewer features present. The blob could be due to simple optical spread of the imaged stimulus because of well known optical effects (e.g., see Cornsweet, 1970) in addition to factors relating to the quite brief display durations. Suppose now that as the blob forms, the observer can perceive roughly whether there is much or little information in the display (e.g., a big blob vs. a little blob), as well as how sharp the image is. We next hypothesize that the observer is able to reset decision criteria at the output level of his feature channels on the basis of this early perceptual information. Alternatively, the observer might be capable of presetting differential criteria for patterns of differential complexity. We suppose that he tends to insert low criteria in the case of big, low-resolution blobs, which usually fol-

low many-featured patterns and are of lower integrity, and to set high criteria with images that possess high resolution. We will see that the major characteristics of the present data follow in consequence. It is hypothesized, in addition, that the integrity (resolution, sharpness) factor is stronger than the "bigness," or complexity, factor.

Consider a typical trial: A stimulus pattern with many features (relative to the total number possible) is shown to the observer. On the average, the image will be a relatively large blob with fairly low integrity. The observer suspects that quite a few features are present but, because of the vagueness of the image, knows he may not be able to clearly detect those present and reject those absent; so both factors dictate that the criteria be decremented, thus making it easier to report features as present. On the other hand, on trials in which a stimulus with fewer features is present, the result will probably be a small blob but with higher integrity. Both factors here direct the observer to raise his criteria.

We conjecture that on trials in which there is conflict between size, or complexity of the early blob form and its integrity, for instance a big blob with high resolution, the resolution determines the setting of the criteria.

This set of hypotheses explains not only the unconditional d' and β results but also those for the microscopic conditional effects. On trials in which the image possesses high integrity, several features will probably be correctly detected and those absent will be correctly rejected. Thus, conditioning on the correct detection of another feature places one within that set of trials in which the integrity of the images was high and thus increases the probability that the given feature was also correctly detected, relative to those trials with low integrity and a decreased likelihood of detecting another feature present in the stimulus. Furthermore, by the same token, conditioning on the correct report of a "real" feature ensures that one is observing data from the subset of trials with higher integrity and therefore higher decision criteria at the featural output stage.

This is, of course, not the only possible line of explanation. For instance, the positive effects of detecting another feature could be due to actual interchannel facilitation, although this possibility does not seem too cogent at the present time. Another potential explanation would hold that attention and/or drift of fixation might provide an artifact such that microanalysis clusters together trials with high integrity due to attention (etc.) rather than, or in addition to, the blob factors mentioned above. Although we cannot yet rule either of these out with total confidence, versions of the latter dynamic that are based on a periodically fluctuating image integrity have not received support from some preliminary auto-

correlation analyses that we have performed on similar data.

The blob model does seem felicitous in that it is consonant with recent theories and experiments on perception in general and multisymbol experiments in particular. Thus, it is in the spirit of the microgenetic information growth cited by Flavell and Draguns (1957) and others, as well as the continuous flow results and theoretical propositions offered in recent years by, for instance, J. E. Hoffman (1975) and Eriksen and Schultz (1979) and, mathematically, by McClelland (1979) and Ashby (1982). Indeed, as Townsend and Ashby have pointed out (1983, chap. 12), the extant continuous-flow models in psychology may be represented as special types of systems that have proven of great value in the physical and biological sciences. The blob conception is also compatible with the idea of different levels of analysis of visual structure (e.g., Kinchla, Solis-Macias, & J. Hoffman, 1983; Navon, 1977) and, potentially, with that of different levels of perceptual invariance (e.g., see Dodwell, 1983; W. C. Hoffman, 1966). The basic conception was first directed toward confusion experiments by Lupker (1979), although not in quite the same fashion as here.

Even if this qualitative model holds in general, it certainly will have to be modified to handle context effects in more complex settings, as in McClelland and Miller (1979) or Prinzmetal (1981). Previous experiments and much theoretical work have also proven the importance of response bias processes, especially when ordinary alphabets or explicit payoff or stimulus frequency manipulations are employed (see, e.g., Townsend, 1971a, 1971b; Townsend & Ashby, 1982; Townsend & Landon, 1982, 1983; cf. Loomis, 1982). Response bias may play a lesser role in the present type of paradigm and, in any case, could not explain the featural correlations obtained in these data.

However these matters are ultimately resolved, it is important to establish the conditions under which such assumptions as those postulated or tested here may hold within varied perceptual environments. Although many experiments have provided evidence that any particular assumption does not hold everywhere, we also need evidence showing where they do hold or, if not, pretty convincing evidence that they are in force nowhere. Not only is this information valuable in its own right as delineating the true characteristics of human information processing, but such assumptions considerably facilitate computation in models of more complex cognitive behavior.

As a single example, the recent investigations by McClelland and Rumelhart (1981) and Rumelhart and McClelland (1982) depend on assumptions about feature processing. They have apparently dropped the high-threshold feature sampling assumption that

was contained in the earlier Rumelhart models (e.g., 1971; Rumelhart & Siple, 1974), which is certainly called for, considering the mounting evidence for the existence of ghost features. It is not clear from their description whether the across-stimulus and across-feature invariance have also been relaxed. They also apparently continue to assume sampling independence, which, as noted, is a tremendously helpful but precarious simplifying assumption. But when such models are tested in word perception or reading experiments one would like to be sure whether their success or failure is due to the higher order principles or due to an incorrect assumption at a lower level of processing. For instance, it remains to be seen whether sampling independence is satisfied in their target data and, if not, whether the violations are sufficient to perturb the "fit" of model to data.

In closing, it is germane to mention the issue of "how much of a difference does it take to make a difference" with regard to model falsification. The present experiment is of reasonable statistical power, relative to the relevant literature. Is it possible that we are picking up on artifacts or epiphenomena that are niggling with respect to the major dynamics of feature processing? Our current assessment, based on the accumulating evidence, is that the deviations found from sampling independence, across-stimulus and across-feature invariance, and high-threshold feature sampling can be sizable and important.

REFERENCES

Ashby, F. G. (1982). Deriving exact predictions from the cascade model. *Psychological Review, 89*, 599-607.

Cornsweet, T. N. (1970). *Visual perception.* New York: Academic Press.

Dodwell, P. C. (1983). The Lie transformation group model of visual perception. *Perception & Psychophysics, 34*, 1-16.

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics, 25*, 249-263.

Estes, W. K. (1972). Interactions of signal and background variables in visual processing. *Perception & Psychophysics, 12*, 278-286.

Flavell, J. A., & Draguns, J. A. (1957). A microgenetic approach to perception and thought. *Psychological Bulletin, 54*, 197-217.

Garner, W. R., & Haun, F. (1978). Letter identification as a function of type of perceptual limitation and type of attribute. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 199-209.

Garner, W. R., & Morton, J. (1969). Perceptual independence: Definitions, models, and experimental paradigms. *Psychological Bulletin, 72*, 233-259.

Geyer, L. H., & DeWald, C. G. (1973). Feature lists and confusion matrices. *Perception & Psychophysics, 14*, 471-482.

Hoffman, J. E. (1975). Hierarchical stages in the processing of visual information. *Perception & Psychophysics, 18*, 348-354.

Hoffman, W. C. (1966). The Lie algebra of visual perception. *Journal of Mathematical Psychology, 3*, 65-98.

Hu, G. G., & Townsend, J. T. (1984). *Feature mechanisms in the visual pattern recognition process.* Manuscript in preparation.

Kinchla, R. A., Solis-Macias, V., & Hoffman, J. (1983).

Attending to different levels of structure in a visual image. *Perception & Psychophysics, 33*, 1-10.

Loomis, J. M. (1982). Analysis of tactile and visual confusion matrices. *Perception & Psychophysics, 31*, 41-52.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. B. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 103-189). New York: Wiley.

Lupker, J. L. (1979). On the nature of perceptual information during letter perception. *Perception & Psychophysics, 25*, 303-312.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 86*, 287-330.

McClelland, J. L., & Miller, J. (1979). Structural factors in figure perception. *Perception & Psychophysics, 26*, 221-229.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. Account of basic findings. *Psychological Review, 88*, 375-407.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*, 353-383.

Prinzmetal, W. (1981). Principles of feature integration in visual perception. *Perception & Psychophysics, 30*, 330-340.

Rumelhart, D. E. (1971). *A multicomponent theory of the confusion among briefly exposed alphabetic characters.* (Tech. Rep. No. 22). San Diego: University of California, Center for Human Information Processing.

Rumelhart, D. E., & McClelland, J. L. (1982). An interaction activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89*, 60-94.

Rumelhart, D. E., & Siple, P. (1974). Process of recognizing tachistoscopically presented words. *Psychological Review, 81*, 99-113.

Schendel, J. D., & Shaw, P. (1976). A test of the generality of the word-context effect. *Perception & Psychophysics, 19*, 383-393.

Townsend, J. T. (1971a). Alphabetic confusion: A test of models for individuals. *Perception & Psychophysics, 9*, 449-454.

Townsend, J. T. (1971b). Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics, 9*, 40-50.

Townsend, J. T., & Ashby, F. G. (1982). Experimental test of contemporary mathematical models of visual letter recognition. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 834-864.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes.* New York: Cambridge University Press.

Townsend, J. T., Hu, G. G., & Ashby, F. G. (1980). A test of feature sampling independence with orthogonal straight lines. *Bulletin of the Psychonomic Society, 15*, 163-168.

Townsend, J. T., Hu, G. G., & Ashby, F. G. (1981). Perceptual sampling of orthogonal straight line features. *Psychological Research, 43*, 259-275.

Townsend, J. T., & Landon, D. E. (1982). An experimental and theoretical investigation of the constant-ratio rule and other models of visual letter confusion. *Journal of Mathematical Psychology, 25*, 119-162.

Townsend, J. T., & Landon, D. E. (1983). Mathematical models of recognition and confusion in psychology. *International Journal of Mathematical Social Sciences, 4*, 25-71.

Wandmacher, J. (1976). Multicomponent theory of perception: Feature extraction and response decision in visual identification. *Psychological Research, 39*, 17-37.

Wandmacher, J., Kammerer, E.-M., & Glowalla, U. (1980). Context dependence in visual feature processing. *Psychological Research, 42*, 335-351.

Watt, W. C. (1980). What is the proper composition of the alphabet? II. Composition. *Ars Semeiotica, 3*, 3-46.

Weisstein, N., & Maguire, W. (1978). Computing the next step: Psychophysical measures of representation and interpre-

tation. In E. M. Riseman & A. R. Hanson (Eds.), *Computer vision systems*. New York: Academic Press.

APPENDIX

The first step in calculating the d′ and β parameters is to determine the probability of a hit (i.e., the probability of sampling a feature present in the stimulus) and the probability of a false alarm (the probability of reporting or sampling a feature not contained in the stimulus) in a particular situation. Consider, as an example, the probability of a hit for the vertical line segment or feature given that ⌐ was presented. In signal detection terms, the vertical feature serves as the signal and the horizontal feature acts as the noise. The probability of a hit would then be the summation of probabilities in the confusion matrix of all cells in which the vertical feature was contained in the response within the ⌐ row. To compute the probability of a false alarm in this case we add the probabilities over all cells in which the signal (i.e., the vertical feature) was contained in the response, but this addition takes place in the row in which the stimulus contained only the noise (i.e., the horizontal feature).

Therefore, to compute the probability of a hit in this instance, we sum cells 2, 6, 7, 8, 12, 13, 15, and 16 of row 7 (⌐) in the confusion matrix. To compute the probability of a false alarm in this instance, we sum the identical cells of row 4 (⌐) in the confusion matrix. Given the probability of a hit and the probability of a false alarm, d′ and β may be computed using the standard signal detection methods.