

A test of visual feature sampling independence with orthogonal straight lines

JAMES T. TOWNSEND, GARY G. HU, and F. GREGORY ASHBY
Purdue University, West Lafayette, Indiana 47907

Well-specified feature detection models of visual character recognition typically assume feature sampling independence; that is, they assume that the detection of one feature is probabilistically independent of the detection of others. Recent results have suggested this assumption may be suspect with letter-like stimuli. The present study utilized very simple stimuli consisting of up to two straight-line segments that were either physically connected or separated by a gap. A strong model that assumed that features are reported if and only if they are sampled together with independence could not be rejected even when the lines were connected.

In all of cognitive psychology, it would be difficult to find a more ubiquitous concept than that of "feature." It has been particularly important in perceptual disciplines due, perhaps, to the hallmark physiological papers by Hubel and Wiesel (1962) on cats and earlier work by Lettvin, Maturana, Pitts, and McCulloch (1961) on frogs, demonstrating neural units associated with specific "features" in the visual stimulus. Although there seems never to have been a completely specified and generally accepted definition of what a feature detector should consist of, there does appear to be some agreement on at least some of the component attributes. One attribute of import in mathematical models that have been developed for letter recognition has been the conception that features are processed or sampled in an all-or-none manner; that is, a feature is either "seen" or completely missed (e.g., see Massaro & Schmuller, 1975). A second commonly assumed property, closely tied to the first, is the idea of probabilistic or stochastic independence; that is, any particular feature is "seen" or "sampled" independently (sampling independence) of any other (e.g., Geyer & DeWald, 1973).¹ This property is an important psychological construct and is usually critical from the standpoint of mathematical modeling. It permits the formulation and testing of a hypothetical set of features within the context of a well-defined mathematical model.

Thus, under sampling independence the probability of two features *a* and *b* being detected is equal to the probability that Feature *a* is detected multiplied by the probability that Feature *b* is detected.

We would like to thank J. T. Berning for help in conducting the experiments. The first and second authors were supported by NSF Grant BNS76-84053, and the third author was supported by Purdue Research Foundation Grant XR0104 when this research was carried out. This research was presented at the annual meeting of the Psychonomic Society, San Antonio, November 1978.

Tests of the base properties of feature sampling are rare, but evidence has arisen that questions the general validity of feature sampling independence (Townsend & Ashby, Note 1). Townsend and Ashby used as stimuli the four letters A, E, F, and H, which were all constructed from equal-length line segments. Observers responded on every trial with both the feature list and the letter that they thought they perceived. These feature list-letter data were used to test five models of the feature sampling process, as well as a number of models of the ensuing letter decision process. The letter decision process was described effectively, but none of the feature sampling models, all of which assumed sampling independence, performed well, thus casting doubt on that postulate.

Wandmacher (1976) employed very simple single-line or single-acute-angle stimuli to test various hypotheses about processing, including sampling independence, which held acceptably. However, no analyses were made of a "strong" model, in which it is assumed that the independently sampled features are essentially reported as sampled. Rather, only a model incorporating an ensuing rather sophisticated decision process was investigated. Also, no blank trials were included, an omission that could perturb the sampling or reporting frequencies of the features.

The present study optimizes the conditions for the possible appearance of sampling independence by employing as operationally defined features two equal-length straight lines positioned at right angles to each other. A strong version of the independence assumption is tested, in which it is hypothesized that in the circumstances of these experiments, a feature (straight line) is reported if and only if it is sampled. The two experiments reported below tested for independence with connected lines as well as with a gap between the lines. Perceptual independence is often thought to depend on retinal separation.

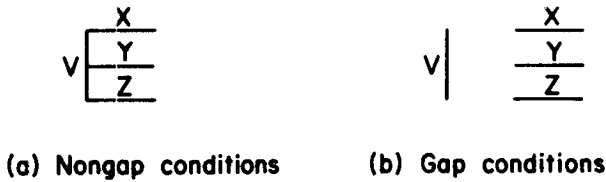


Figure 1. Feature representation of the stimuli used in Experiment 1.

EXPERIMENT 1

Method

Subjects. The observers were four Purdue upper division majors in psychology with 20/20 vision.

Apparatus. A Gerbrands two-field tachistoscope (Model T-2b) was used to present the stimuli. The operational stimulus features were given the names V, X, Y, and Z, as shown in Figure 1, with V = vertical, X = top horizontal, Y = middle horizontal, and Z = bottom horizontal. G indicates a gap between the two lines in the stimulus. The stimuli were VXG, VYG, VZG, VX, VY, VZ, V, X, Y, Z, and blank; these were arranged into three sets. Each set was composed of one of the three horizontal lines, the vertical line, the blank, and both the vertical and horizontal lines together, with or without the gap. There were therefore five stimuli in each set. All stimuli were constructed from equal-length line segments; both the gap size and line length were .45 cm (25 min visual angle). The gap stimuli were produced by moving the horizontal line to the right from the vertical line. A prestimulus fixation field was described 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 one side subtended an angle of about 2 deg at the observer's eye. Luminance was maintained at about 8 fL. Responses were given verbally and were recorded by the experimenter on recording sheets.

Procedure. The different horizontal lines were presented in separate counterbalanced blocks, but the gap and nongap conditions were within blocks. The observers had 5 days practice for a total of 5 h. Before the experiment began, stimulus duration was set for each observer individually so the probability of being correct was around .45. Twenty trials of practice preceded each of 15 experimental sessions. There were three blocks of 80 trials per day, one associated with each horizontal line. Observers were instructed to report only the features they thought they had seen and not to report any gaps. Thus, for example, the correct response to both Stimuli VX and VXG was VX. Trials were counterbalanced to equate frequency of response. Thus in Block 1 over the 15 experimental days, each observer was presented Stimuli VX and VXG 150 times each. The observers were presented each of the other stimuli in the block 300 times. The blocks were suitably counterbalanced across sessions.

Results and Discussion

The focus of the present study was on the sampling independence assumption. Other aspects of the data will be considered elsewhere. Toward this end, the most pertinent data came from trials on which the stimulus was VX, VY, or VZ. This is because sampling independence assumption states that the feature sampling probabilities are multiplicative and thus $P(VX) = P(X) \times P(V)$, $P(VY) = P(V) \times P(Y)$, and $P(VZ) = P(Z) \times P(V)$. Under this assumption we can obtain a maximum likelihood estimate of the marginal probability of sampling, say, Feature X when Stimulus VX is presented by $P_{VX}(VX|VX) + P_{VX}(X|VX)$, where $P_{VX}(X|VX)$ is the observed probability of Response X given Stimulus VX in Condition VX. Similarly, $P_{VX}(V) = P_{VX}(V|VX) + P_{VX}(VX|VX)$, and so on. Thus, under independence, the maximum likelihood estimate of the probability of sampling VX when the stimulus is VX is $P_{VX}(VX|VX) = [P_{VX}(X|VX) + P_{VX}(VX|VX)] \times [P_{VX}(V|VX) + P_{VX}(VX|VX)]$. Both this prediction and the observed value for each condition are shown in Table 1.

This fact allows the construction of a likelihood ratio test of the null hypothesis of independence. The results of this analysis are shown in Table 1. It can be seen that about half of these tests indicated a dependence in the data and about half supported independence. The surprising thing about these results, however, is that almost all the suggested dependencies come from the gap condition. Only 3 of 12 tests on the nongap data supported dependence, whereas 10 of 12 tests on the gap data did.

The dependence found in the gap conditions is puzzling. To verify that the result was not an artifact of our test statistic, we performed a chi-square goodness-of-fit test between the independence model and the data using the iterative program STEPIT (Chandler, 1965) to estimate the model parameters. The difference between this method and the likelihood ratio test is that the former is not constrained to use maximum likelihood estimates as is the likelihood ratio test. Instead, STEPIT selects those parameter values that yield the smallest value of the chi-square statistic. The results were virtually identical: independence in the nongap condition and dependence in the gap condition. The closeness of the two analyses also reaffirms the optimality of the maximum likelihood estimates.

Let us consider in more detail the stimuli used in,

Table 1
Results of Statistical Analyses for Data of Experiment 1

Stimulus	Observer 1			Observer 2			Observer 3			Observer 4		
	O	P	LRT	O	P	LRT	O	P	LRT	O	P	LRT
VX	.127	.152	1.67	.174	.154	.98	.187	.239	2.05	.260	.252	.19
VXG	.393	.291	26.00*	.347	.304	6.39*	.373	.315	9.19*	.600	.529	20.67*
VY	.153	.249	23.85*	.320	.344	1.50	.407	.449	6.14*	.260	.261	.62
VYG	.180	.178	.02	.213	.152	10.26*	.313	.281	2.65	.487	.402	15.71*
VZ	.164	.256	23.14*	.473	.448	1.96	.507	.500	.17	.413	.402	.47
VZG	.247	.187	10.05*	.360	.262	30.57*	.367	.260	28.42*	.440	.358	17.79*

Note—O = observed probability of Response R_i (e.g., VX, VY, VZ) given stimulus; P = predicted probability of Response R_i from maximum likelihood estimate; LRT = likelihood ratio test. * $p < .05$.

say, Block 1 of the experiment. Assume that the observers really do process the stimulus features in an independent fashion. Now consider what might happen on trials when the stimulus is VX. With some probability, the observer will not see Feature V and will see Feature X and hence will respond "X." Suppose, on the other hand, the stimulus is VXG and again the observer does not see Feature V but does see Feature X. If he also sees that Feature X is displaced to the right to a certain extent, the observer may realize that Feature V must also have been presented, since VXG is the only stimulus configuration in which X is displaced to the right. Thus, in these circumstances, the observer may respond "VX" even though he never saw the Feature V. This strategy would destroy independence, since there will tend to be more VX responses and fewer X responses than independence predicts. This is precisely what was found in the data. In all 12 of the gap conditions, it was found that $P(Vi|ViG) > [P(Vi|ViG) + P(V|ViG)] \times [P(Vi|ViG) + P(i|ViG)]$ (where $i = X, Y, \text{ or } Z$), whereas independence predicts an equality.

To test this hypothesis and gain further evidence on the independence question, a new set of stimulus conditions was designed and employed in Experiment 2. First, the gap and nongap conditions were run in separate blocks, and second, both the horizontal and the vertical lines were displaced to create the gap stimulus (in Experiment 1, only the horizontal line was perturbed). Because the three horizontal lines yield similar results, only the top line (X) was used in Experiment 2. There were several ancillary conditions, involving subsets of the various stimuli, run together (but in separate blocks) with the major conditions of Experiment 2. These are irrelevant to the questions asked here and will be omitted from the Method and Discussion sections.

EXPERIMENT 2

Method

Subjects. The observers were four Purdue upper division majors in psychology with 20/20 vision, none of whom participated in Experiment 1.

Apparatus and Stimuli. The apparatus used was the same as in Experiment 1. The stimulus features were the V and X line segments of Experiment 1. Two stimulus sets were constructed, with one being composed of (1) vertical lines, (2) horizontal lines, (3) both together without gap, and (4) blank. The second set was identical to the first except for the existence of a gap in the combined stimuli. The two stimulus sets were presented in

separate counterbalanced blocks. The gap size was, as in Experiment 1, approximately 25 min visual angle. When 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 nongap blocks. Similarly, Stimulus Feature X was displaced an equal distance to the right in all gap blocks. Preexperimental practice and calibration was the same as in Experiment 1. Each block was run once a day for 15 days. Within each block, each stimulus was presented 10 times. There were thus 150 trials per data point per observer.

Results and Discussion

The same analyses performed in Experiment 1 were carried out on Experiment 2 data. The results of the likelihood ratio test are given in Table 2, along with the observed and predicted conditional probabilities of responding "VX" when the stimulus was VX or VXG. Again, the likelihood ratio test statistic ($-2 \log L$, where L is the likelihood ratio) has an asymptotic chi-square distribution with 1 degree of freedom. None of the values are significant at the .10 level. Thus the assumption of independence cannot be rejected in any of the conditions. A minimum chi-square test yielded identical results. It appears that our hypothesis about the cause of the dependence in the gap conditions of Experiment 1 was correct.

These results appear at first to contradict the aforementioned results where support for sampling dependence was found in a letter recognition task (Townsend & Ashby, Note 1). One hypothesis that resolves the apparent conflict is that dependence is caused by the increased prominence of decisional influences when more complex and familiar letter-like stimuli are used. For instance, in letter recognition tasks, some combinations of features may appear "impossible" to the observer in the sense that they are contained in no single stimulus, and thus they may be less likely to end up in the observer's internal list of sampled features than strict independence predicts. This "filtering out" of impossible feature combinations could occur almost anywhere in the system. For instance, the system might allocate more capacity to combinations of peripheral complex feature detectors (set to detect a combination of simple features) that are consonant with the known set of stimuli. On the other hand, the filtering might come much later and involve more traditional decisional influences. The latter hypothesis seems more plausible, but no evidence currently exists to test between them.

In the present study, no such impossible feature sets

Table 2
Results of Likelihood Ratio Tests for Experiment 2

Stimulus	Observer 1			Observer 2			Observer 3			Observer 4		
	O	P	LRT	O	P	LRT	O	P	LRT	O	P	LRT
VX	.195	.191	.03	.567	.556	.59	.327	.301	1.75	.260	.240	1.11
VXG	.165	.187	1.46	.515	.538	2.33	.263	.277	.44	.187	.188	.04

Note—O = observed probability of Response VX; P = predicted probability of Response VX from maximum likelihood estimate; LRT = likelihood ratio test.

exist, yet dependence was still found in the gap conditions of Experiment 1. Again, a decision influence was supported by the analysis. However, a hypothesis that more spatial interactions occur in more complex stimuli that cause dependencies cannot be presently ruled out. Further, Wandmacher's (1976) results are compatible with the idea of independent sampling followed by a decisional stage.

We tentatively conclude that the assumption of sampling independence may provide a reasonable first approximation under sufficiently circumscribed experimental conditions. We do not make the impossible claim to have proven the null hypothesis that no dependence exists.

Future studies are planned to investigate the experimental parameters affecting dependence. One study will employ synthetic characters that are more complex but contain all possible feature combinations. Thus, no sampled set of features would be impossible, providing evidence for the above hypothesized decisional influence.

REFERENCE NOTE

1. Townsend, J. T., & Ashby, F. G. *Testing contemporary models of letter recognition*. Paper presented at the annual meetings of the Midwestern Psychological Association, Chicago, 1976.

REFERENCES

- CHANDLER, J. P. *Subroutine STEPIT: An algorithm that finds the values of the parameters which minimize a given continuous function*. Bloomington, Ind: Quantum Chemistry Program Exchange, Indiana University, 1965.
- GEYER, L. H., & DEWALD, C. G. Feature lists and confusion matrices. *Perception & Psychophysics*, 1973, **14**, 471-482.
- HUBEL, D. H., & WIESEL, T. N. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology*, 1962, **160**, 106-154.
- LETTVIN, J. Y., MATURANA, H. R., PITTS, W. H., & McCULLOCH, W. S. Two remarks on the visual system of the frog. In W. A. Rosenblith (Ed.), *Sensory communication*. New York: Wiley, 1961.
- MASSARO, D. W., & SCHMULLER, J. Visual features, preperceptual storage, and processing time in reading. In D. W. Massaro (Ed.), *Understanding language*. New York: Academic Press, 1975.
- WANDMACHER, J. Multicomponent theory of perception: Feature extraction and response decision in visual identification. *Psychological Research*, 1976, **39**, 17-37.

NOTE

1. Because the term "feature extraction" often carries connotations of questionable peripheral physiology, we prefer the more neutral term "feature sampling."

(Received for publication January 7, 1980.)