

The statistical analysis of concurrent detection ratings

THOMAS D. WICKENS and LYNN A. OLZAK
University of California, Los Angeles, California

The concurrent detection task is a powerful method for assessing interactions in the processing of two sensory signals. On each trial, a stimulus is presented that is composed of one, both, or neither signal, and the observer makes a detection rating for each stimulus. A classical bivariate signal-detection analysis applies to these data, but is limited by its inability to differentiate certain types of sensory interactions from more cognitive components, and by the lack of an associated testing procedure. The present paper presents an alternative analysis, based on the contingency table of sensory ratings. Six classes of effect can be distinguished and tested: (1) simple response bias, (2) detection of the two signals, (3) interference of each signal on the response to the other signal, (4) sensory and response correlation, (5) bivariate response biases, and (6) higher order association. Complete computational detail is provided.

The interaction among different sensory systems, or among mechanisms that subserve a single sensory modality, is a critical aspect of perceptual function. How best to study these interactions is a challenging problem, as evidenced by the multitude of approaches, analyses, and interpretive arguments that have flourished over the years. Classical approaches to the study of interactions, such as masking, adaptation, and subthreshold summation, rely on an analysis of how the presence or absence of one signal affects the ability to detect or identify a second, target signal. Such experiments indicate whether an interaction exists and provide information about its direction (excitatory or inhibitory). They have been essential in developing multiple-mechanism models of visual, auditory, and cutaneous sensory processes (for reviews, see Olzak & Thomas, 1986; Scharf & Buus, 1986; Sherrick & Cholewiak, 1986). Recent variations of these classical techniques, such as profile analysis (Green, 1983; Green, Mason, & Kidd, 1984) and simultaneous detection and identification (Olzak, 1985; Thomas, 1985), have been successful in isolating interactions among widely separated mechanisms in the auditory and visual systems. However, these paradigms provide little detailed information about the source of the interaction. They do not distinguish among various sensory candidates, or between the purely sensory and more cognitive components.

In this paper, we focus on the statistical analysis and interpretation of a more informative psychophysical technique: the *concurrent rating task*. This paradigm, a natural extension of the signal-detection-theory rating task to mul-

iple dimensions, yields information about both the intended signal and its competition. Although the utility of a signal-detection analysis in isolating various types of interactions has been previously demonstrated (e.g., Hirsch, Hylton, & Graham, 1982; Olzak, 1985), it has serious limitations as usually applied and does not lead to an unambiguous interpretation of all results (Ashby, 1988; Ashby & Townsend, 1986; Klein, 1985; Olzak & Wickens, 1983; Olzak, Wickens, & Gouled-Smith, 1985). Treatment of the data as a frequency table to which log-linear models are fit avoids some of these limitations (Olzak & Wickens, 1983), but does not fully disentangle certain types of sensory interactions from more cognitive components such as response bias.

In this paper, we approach the analysis of concurrent rating data through a logical sequence of hypothesis-testing steps designed to resolve many of the ambiguities associated with earlier approaches. For each hypothesis, appropriate models and associated statistical tests are described and illustrated by example. Our example is drawn from research on the visual system, but our procedures apply equally well to any modality. The paper is organized in four sections: (1) a description of the task itself and a concrete example, (2) procedures designed to test hypotheses of independence, (3) procedures designed to describe the nature of complex associations isolated in the second step, and (4) a mathematical and statistical appendix that describes the details of our procedures. The first three sections introduce the problems and provide a step-by-step protocol for resolving them. A reader who is unfamiliar with some portions of the statistical analysis, or who wishes to duplicate it exactly, should refer to the Appendix as needed.

THE CONCURRENT RATING TASK

As noted above, the simplest way to study signal interaction is to examine how the detection or identification

This work was supported in part by U.S. Public Health Service Grant EY00360 from the National Eye Institute to James P. Thomas, and by a UCLA Faculty Research Grant to the first author. We thank Jim Thomas for this support and for much discussion, and Greg Ashby, Eric Holman, and an anonymous reviewer for comments on earlier drafts. Requests for reprints and information about the computer programs may be addressed to Thomas D. Wickens, Department of Psychology, University of California, Los Angeles, CA 90024.

of one signal is influenced by the presence or absence of another signal. Unfortunately, the interpretation of this simple paradigm is ambiguous. Because there is no way to assess the magnitude of the response to the competing stimulus, one cannot separate the effects of the stimulus from those of the response. In short, the signal is known, but the way it is perceived is not.

A natural way to start to sort out these effects is to collect data about both the intended signal and its competition. The logical experiment uses a *concurrent rating task*. An experimenter presents a stimulus that is composed of both a test signal and an interfering signal, and obtains detection ratings from the subject for both of these signals. Specifically, the two signals form four types of stimuli, according to whether each is present or absent. The subject makes a pair of judgments that express his or her certainty that each of the signals has occurred. This design has a nice compactness: it is symmetric with respect to the two signals, so the experimenter can simultaneously look at the influence of the first signal on the second and of the second signal on the first.

The terminology we use here bears emphasis. Throughout this paper, the word *signal* refers to the pure components that are to be detected, while the word *stimulus* refers to whatever combination of signals is formed into a single presented entity. For example, in the data we will discuss below, the signals are spatially modulated sinusoidal gratings with different frequencies. These two signals (if present) are superimposed into a single stimulus. In contrast to the overt decomposition of a stimulus into signals, it is often useful to think of hypothetical mechanisms within the observer that respond, with greater or lesser selectivity, to the separate components of the stimulus and upon which a response is based. We use the word *channel* to refer to these hypothetical mechanisms. Very often these channels are identified with the particular signals, although it is not necessary to do so.

The response collected from this two-signal concurrent detection task is composed of two parts, each pertaining to one of the signals. In the data that we discuss, these judgments are recorded as a pair of *k*-level ratings, each of which categorizes one of the signals on a scale that ranges from definitely absent to definitely present. This composite response can be tallied as one cell of a $k \times k$ table. The complete data from such an experiment consist of four $k \times k$ tables, or of one $2 \times 2 \times k \times k$ table, corresponding to the four stimulus conditions.

In the following discussion, we consider data from an experiment by Olzak and Kramer (1984) on the detection of compound sinusoid spatial gratings.¹ The stimuli in this experiment were spatial modulations of a constant background luminance. The modulations were sinusoidal gratings of high and low spatial frequency (12 and 3 cycles per degree, respectively). The composite stimuli were presented to the subject for 100 msec. Ratings were obtained for both components as a $k=6$ level confidence rating. Each of the four stimulus conditions was rated 350 times. Table 1 shows the data from one subject, which

Table 1
An Example of Concurrent Detection Data

Low-Frequency Signal		High-Frequency Signal X											
		Absent						Present					
		Y	1	2	3	4	5	6	1	2	3	4	5
Absent	1	44	4	9	7	6	7	7	4	5	5	14	69
	2	13	30	20	8	14	7	5	7	13	15	38	37
	3	9	23	17	17	3	0	6	7	8	10	10	15
	4	16	17	10	20	2	2	4	12	5	13	6	14
	5	5	4	9	10	4	0	2	3	1	1	3	5
	6	3	3	0	1	4	1	0	0	1	1	1	3
Present	1	8	2	2	1	0	4	4	1	2	0	4	37
	2	5	5	5	5	5	3	0	4	0	1	8	25
	3	8	10	7	4	1	1	1	3	3	7	8	15
	4	12	17	15	13	2	2	4	4	8	17	12	21
	5	12	17	19	18	10	4	3	12	8	11	20	20
	6	31	29	25	24	12	12	11	8	12	11	12	33

are used as an example throughout this paper. As noted, this table constitutes a four-dimensional array, the dimensions of which are symbolized by the signal dimensions *H* and *L* (for "high" and "low") and the response dimensions *X* and *Y*. The *X* response is appropriate for the *H* signal and the *Y* response for the *L* signal. Entries in Table 1 indicate the number of times that each response combination was made; for example, on 13 of the 350 trials on which a stimulus with no signal components was presented, the subject responded with a rating of $X=1$ for the high-frequency signal and $Y=2$ for the low-frequency signal. These frequencies² are indexed by four subscripts and are denoted here by f_{hlyx} , where the subscripts *h* and *l* take the two values *a* and *p* (for "absent" and "present"), and the subscripts *x* and *y* take values from 1 through 6. Capital letters refer to the factors or dimensions of the table in general; lowercase letters refer to particular values.

Influences on the Responses

Before turning to an analysis of Table 1, it is helpful to make clear the general view of signal detection that we are using. Without some model of the detection process, the reader is apt to become confused as to where the interactions lie. A plausible (if somewhat naive) view is shown in Figure 1. The two signals, *H* and *L*, are combined to form a joint stimulus. These components need not, of course, be high- and low-frequency signals, but we will use these letters to be consistent with our example. The subject is presented with this stimulus, which results in a joint percept. The perceptual analysis may take

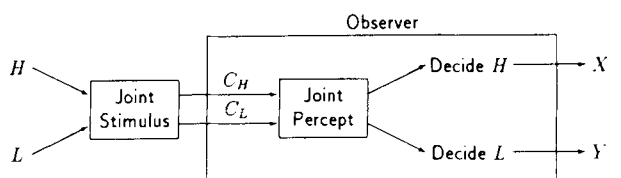


Figure 1. A schematic model of the signal-detection process.

place through a pair of channels labeled C_H and C_L , roughly tuned to the frequencies of H and L . It is not necessary for these channels to be distinct from each other; they surely would not be distinct if the signals were close together. In any case, either or both of the signals may be detected in the joint percept. The conventional statistical-decision model of signal-detection theory is adequate to describe what goes on at this stage. Finally, two responses, X and Y , are made, one for each signal.

There are many reasons for the frequencies in Table 1 to take the values that they do. Before attempting a quantitative analysis of the concurrent detection data, it is essential to define the potential sources of influence. Starting with the conceptual framework in Figure 1, at least six sources can be identified:

1. *Response bias*. The subject may prefer certain numerical categories to others. For example, a subject may use Category 4 frequently and Category 5 rarely.

2. *Detection*. The responses are almost certainly influenced by their proper signal. The presence of the high signal ($H=p$) produces higher X ratings than does its absence ($H=a$). Likewise, L and Y are associated.

3. *Sensory interference*. The responses may be influenced by the presence or absence of their opposite signal, so that H affects Y and L affects X . Generally, one expects these influences to take the form of simple sensory excitation or inhibition. These effects act at the sensory end of Figure 1, for example as some sort of "cross-talk" between the detection channels C_H and C_L . Ashby and Townsend (1986) referred to the lack of sensory interference as *perceptual separability*.

4. *Correlated sensory noise*. An interaction between the responses may be produced by noise that is introduced during the formation of the joint percept and that simultaneously influences both channels. Such noise affects responses X and Y simultaneously. The result is not the same as excitation or inhibition by the opposite signal, since it creates association within a single stimulus condition. Ashby and Townsend (1986) referred to the lack of correlated sensory noise as *perceptual independence*.

5. *Bivariate numerical biases*. Just as the subject may prefer one numeric response to another, certain combinations of responses may be preferred or avoided. Like the correlated sensory noise, these numerical or response strategies create bivariate association patterns involving both X and Y . For example, a subject may prefer (or avoid) using the same number for both responses, or may prefer (or avoid) making high certainty ratings (Categories 1 or 6) on both responses. Similarly, a subject who fails altogether to pay attention to the stimulus on occasional trials will show an excess of 1,1 responses.

6. *Higher order associations*. Associations that involve more than two factors are potentially present, although they are much more difficult to understand. Such an association would be created if, for example, the magnitude of the correlated portion of the noise depended on the level of one of the signals. Present perceptual theories make few predictions about such effects. They should be considered only if simpler mechanisms are inadequate to

explain the data. However, even if one does not understand these effects, a method is needed to recognize them.

These six types of effect are confounded with each other and cannot be unambiguously sorted without further assumptions. The separation of bivariate numerical biases and their correlated sensory responses illustrates the problem. If these effects are allowed to take a completely general form, one can never decide whether a particular bivariate pattern should be attributed to one factor or the other. Nevertheless, certain forms of association are more probable than others. As we argue in this paper, a few assumptions about the plausible forms of the effects allow a substantial degree of separation to be made.

We have listed the six effects above in approximately the order that we felt they should be introduced into a description of the signal-detection process. Two factors determined this ordering. First, an explanatory conservatism dictated that low-order effects (e.g., response bias) should be introduced before including multifactor effects in our explanations. Second, certain explanations that involve a given number of factors are simpler than others and should be considered first. For example, a description based on correlated response channels (Source 4) should not be used if the same data can be explained equally well using only signal-response effects (Sources 1, 2, and 3). Thus, we considered correlated response channels only as an addition to a model that already included response bias, detection, and sensory interference. We used these principles to direct our choices of tests and explanations below.

ANALYSIS OF INDEPENDENCE AND CONDITIONAL INDEPENDENCE

In previous work (Olzak & Wickens, 1983; Olzak, Wickens, & Gouled-Smith, 1985; Wickens, 1989), we have argued for the utility of fitting concurrent rating data with linear models for the logarithm of the frequency (i.e., log-linear models). These models enable hypotheses about specific relationships among the factors to be tested. We will not reiterate our discussion of model fitting here, but instead approach the problem with a hypothesis-testing logic, using null hypotheses of conditional independence. Although we use models similar to those we have described before, their interpretation is different.

Tests of Association

In this section, we examine the six categories of effect described above and describe how they are identified and expressed in a set of data.

Response bias. We will discuss response-bias effects the least. They are manifested in the marginal frequencies with which the categories are used, that is, in the marginal distributions of the responses X and Y . These distributions depend primarily on the attractiveness of the numerical categories. Thus, the subject may choose to avoid or to rely on a particular response category, giving it a lower or higher frequency than might otherwise occur. The distribution is also affected by signal strength,

since a strong signal would produce responses that would fall into the most extreme categories.

The distribution of category usages can be obtained by collapsing the data to the one-dimensional distributions of X or Y responses. However, these distributions are less informative about category usage than one might expect, since they are dictated, in part, by the detectability of the signals. With highly detectable signals, few intermediate responses appear. Moreover, an evaluation of category use is not independent of the model that is used to describe the data. To avoid letting our analysis of association become confounded with assumptions about response preferences, we include free parameters in our models that match their predictions to the observed marginal distributions.

Detection. Although the data in Table 1 form a four-way contingency table, many hypotheses about the relationships among the factors are tested in simpler tables. One such hypothesis is detection of the signal. This detection is the subject's task, and it is the first association that one should look for in the data. For example, Signal H should be associated with Response X . Because this association has explanatory primacy over the effects that relate X to the factors L or Y , the latter two dimensions can be eliminated from the table. This simplification treats any detection-like association as actually caused by detection, rather than as secondary to some interaction with the other signal. With this assumption, the association between H and X is examined statistically by collapsing Table 1 into a two-factor HX table as shown in Table 2.

In the conventional analysis of Table 2, the frequencies are converted to cumulative proportions along the rows and a receiver-operating-characteristic curve is constructed. The detectability of the H signal is then measured by calculating either d' or the area under this curve. Another approach, albeit a somewhat weaker one, is to ask if there is any association between the H and the X classifications. In elementary statistics, this hypothesis is tested using a chi-square test of homogeneity. The null hypothesis for this test is of no association between H and X . We denote this lack of relationship by the expression $X \parallel H$, which can be read "X is unrelated to H." This hypothesis is tested in the usual way, using either the Pearson goodness-of-fit statistic or the likelihood-ratio statistic (see Appendix A-1, Equations 2 and 3). For the data in Table 2, homogeneity is easily rejected with a likelihood-ratio test statistic of $G^2 = 381.72$. A similar conclusion holds for L and Y , with $G^2(L \parallel Y) = 413.03$. Obviously, the signals are related to their appropriate responses.

Table 2

Marginal Distribution of H and X Based on the Data in Table 1

H	X					
	1	2	3	4	5	6
a	166	161	138	128	63	43
p	47	65	66	92	136	294

These tests are relatively weak in that they require no assumptions about the relationship of the responses to the subjective stimulation, even that it be monotone. It would be possible for these tests to identify patterns of responding that are psychologically unreasonable—for example, a pattern in which Response 5 is used more often in the absence of the stimulus than Responses 4 and 6 are used in the presence of the stimulus. In fact, an examination of association parameters, as described in the more interesting case of sensory interference below, shows that the subject is using the response categories in a sensible manner, with high-numbered responses associated with the presence of the signal and low-numbered responses associated with its absence.

Sensory interference. The other types of relationships postulated above are also tested by null hypotheses of unrelatedness. However, these effects are not bivariate and must be tested in tables that involve three or more factors. Fortunately, the unrelatedness hypotheses also have multidimensional forms.

After direct signal association, the next effect that should be examined is sensory interference. For example, one can look at the effect of the low-frequency Signal L on Response X that is appropriate to the H signal. In examining this relationship, the context of the appropriate signal needs to be taken into consideration. Because the X response is heavily influenced by the H signal, it is inappropriate to test for LX relationships by investigating the hypothesis $L \parallel X$ in a two-dimensional table (see Appendix A-1). Instead, allowances for the effects of H are made by working in the three-factor HLX table and testing the hypothesis $L \parallel X | H$. When applied to the data in Table 1, these effects are substantially smaller than those that involve the HX and the LY effects. The hypothesis $L \parallel X | H$ is retained—compare $G^2_{10}(L \parallel X | H) = 4.49$ to a 5% critical value of 18.31. The other test shows a modest effect: $G^2_{10}(H \parallel Y | L) = 45.26$.

A significance test demonstrates the existence of an association, but does not indicate its form. When $H \parallel Y | L$ is rejected, as it is for Table 1, the HY relationship should be examined. Two procedures are helpful here, the examination of residuals and the estimation of the parameters of the log-linear model (see Appendixes A-2 and A-3). Looking at the residuals is the most direct procedure. Table 3 shows the standardized residuals from the model $H \parallel Y | L$. Cells in which there is a failure of conditional independence have deviations with relatively large absolute values. The pattern of values in Table 3 is quite clear. When the H signal is present, low-numbered responses are overrepresented (the standardized deviates are positive); when it is absent, they are less frequent. The effects are similar for both levels of L , and indicate an inhibitory effect of H on Y .

The second approach to interpreting the effect is to fit a model that expresses dependence. In log-linear form, this model includes both the terms of $H \parallel Y | L$ and an additional association term $\lambda_{HY(hy)}$. This model is denoted $[HL][HY][LY]$ and has the log-linear form

Table 3
Association of H and Y for the Data in Table 1

		Y					
		1	2	3	4	5	6
Deviations from the hypothesis $H \perp\!\!\!\perp Y L$							
$L = a$							
$H = a$		-1.40	-1.12	0.83	0.85	1.76	1.00
$H = p$		1.40	1.12	-0.83	-0.85	-1.76	-1.00
$L = p$							
$H = a$		-2.72	-0.87	-0.51	-0.31	0.34	2.19
$H = p$		2.72	0.87	0.51	0.31	-0.34	-2.19
Parameters $\lambda_{HY(hy)}$ in the model $[HL][HY][LY]$							
$H = a$		-0.30	-0.18	-0.00	0.03	0.16	0.30
$H = p$		0.30	0.18	0.00	-0.03	-0.16	-0.30

$\log \mu_{hly}$

$$= \lambda + \lambda_{H(h)} + \lambda_{L(l)} + \lambda_{HL(hl)} + \lambda_{Y(y)} + \lambda_{LY(l_y)} + \lambda_{HY(hy)}$$

This model fits very well, having $G^2_5 = 3.25$. The bottom portion of Table 3 shows estimates of the parameters $\lambda_{HY(hy)}$ under this model. Their values substantiate the analysis of the residuals; the association is such that it creates more high-valued Y responses when H is absent than when H is present. For example, $\hat{\lambda}_{HY(a1)} = -0.30$, meaning that $\hat{\mu}_{a1}$ is smaller and there are fewer $Y=1$ responses when the high-frequency signal is absent than would otherwise be expected. Again, inhibition of Y by H is implicated. In interpreting Table 3, one should recognize that the antisymmetric form of $\hat{\lambda}_{HY(hy)}$ is an artifact resulting from the standardization that makes this parameter sum to 0 over either of its subscripts. One cannot conclude, for example, that the absence of H actually stimulates a Y response.

Response association. The final set of independence hypotheses concern the relationship between the two responses, X and Y . These associative effects are the most difficult to untangle, as they subsume both correlated noise and numerical bias, as well as more complicated effects. Once again, these relationships are studied in conditional form. It is quite obvious in Table 1 that the form of the XY relationship is influenced by the signals. For example, when $H=L=a$, the bulk of the responses have low numbers, whereas when $H=L=p$, most responses are high-numbered. Pooling these disparate forms into an XY table may attenuate existing effects or induce spurious ones. Our analysis of the detection process gives some theoretical justification for conditionalization as well. In the list of effects, the associations between the signals and the responses precede those that imply pure XY association. To express this ordering, tests for the independence of X and Y should be constructed so that any association of these factors to the H and L classifications is excluded. The appropriate null hypothesis is $X \perp\!\!\!\perp Y | HL$. This hypothesis is readily rejected, with the statistic $G^2_{100} = 306.45$. Some form of response-response association is present.

With 100 degrees of freedom, there are many ways that $X \perp\!\!\!\perp Y | HL$ can be violated. On the basis of this test of independence, the possible sources of the effect cannot be separated. Numerical bias, correlated noise in the channels, the direct influence of one sensory event on the other response, and other effects may be involved. Separating these effects requires further assumptions about their nature and is discussed in a later section. However, it is possible to rule out the higher order associations at this stage. When a bivariate association term, $\lambda_{XY(xy)}$, is added to the conditional independence model $X \perp\!\!\!\perp Y | HL$, giving the log-linear model $[HLX][HLY][XY]$, the fit is fairly good, $G^2_5 = 92.62$. It appears that this two-dimensional term can explain most of the XY effect, and that higher order associations involving X , Y , and one or more of the signal levels are unnecessary.

An Additive Decomposition of the Effects

A useful characteristic of the conditional-independence tests is that, if appropriately chosen, they form an additive decomposition of the failure of simple models to fit the data. The hypotheses tested in the last section are part of an orthogonal decomposition of the relationships among the four factors in the data (i.e., among H , L , X , and Y). To see this representation, it is necessary to introduce two further models for Table 1. Neither of these models represents a conditional-independence hypothesis, although both represent the intersection of several such hypotheses. Both can be written in log-linear form.

The simplest of the two models represents the complete independence of the responses from each other and from the stimulus, $[HL][X][Y]$. This model describes the performance of a blind observer who generates two random, unrelated responses. Rejection of this model implies that there is some association between the stimulus and the response, or between the two responses. The second model, $[HL][HX][LY]$, adds associations between the signals and their proper responses to the blind-observer model. An observer whose responses are influenced only by the signals that are called for by the task would fit this model. Rejection of this model implies the presence of associations other than those explicitly demanded by the task.

The relationship between tests of these models and tests of the conditional-independence hypotheses are shown in Table 4. Essentially, this table shows the decomposition of the total effect into a series of separate tests. The random-response model $[HL][X][Y]$ represents a composite hypothesis that combines five types of independence or conditional independence. The value of G^2 by which this model is rejected is partitioned in the next three lines into components associated with the two proper responses (indicated by the hypotheses $H \perp\!\!\!\perp X$ and $L \perp\!\!\!\perp Y$) and a test of the hypothesis that these are the only relationships observed. In the final three lines, the latter statistic is further divided into parts associated with the three extra-instructional associations. Throughout, both G^2 and the degrees of freedom are additive.

Table 4
Tests of Independence and Conditional-Independence Hypotheses,
and Information-Transmission Statistics for the Data in Table 1

Hypothesis or Model	G^2	df	T
Random responses, $[HL][X][Y]$	1150.95	130	
$H \parallel X$	381.72	5	0.197
$L \parallel Y$	413.03	5	0.213
Proper responses, $[HL][HX][HY]$	356.20	120	
$H \parallel Y L$	45.26	10	0.023
$L \parallel X H$	4.49	10	0.002
$X \parallel Y HL$	306.45	100	0.158

An index of the size of the effects in Table 4 is also useful. The value of G^2 is unsatisfactory here, as it is influenced by the size of the sample. For the tests of independence, a natural choice for this measure is the *information transmission* between factors (see Appendix A-4). This quantity is essentially a rescaling of G^2/N , but can be interpreted as the number of bits of information in one classification that is carried over to the other. This metric allows for comparison of the sizes of the effects. In this rather difficult detection task, each signal transmitted only about 0.2 bits of information to the response (out of a potential maximum of 1 bit). The residual interresponse effect is substantial; the XY transmission of 0.158 bits is about 80% of the transmission from either stimulus to its proper response.

The decomposition of Table 4 is not unique but is determined by the ordering of the effects. For example, it places the direct association of H and X before the conditional association of L and X given H , or of X and Y given H and L . One could order the factors differently and obtain a different decomposition. As in most hierarchical series of tests in statistics, the structure that is revealed is partially driven by theory. We believe that our organization of the effects is sensible, but we recognize that a researcher who approaches the data with a different view of the perceptual process might choose to order the associations differently and to create a different partition.

To summarize the results of these tests, Table 4 shows that (1) the signals are substantially detected, (2) detection is not all that is going on, (3) there is no tendency for L to influence the X response, but there is some influence of H on Y , and (4) there is substantial association between the responses. In the next section of this paper we explore this XY association.

ANALYZING RESPONSE-RESPONSE ASSOCIATIONS

At this point we have gone as far as we can with an analysis based on conditional-independence concepts. The statistic $G^2(X \parallel Y|HL) = 306.45$ clearly indicates the presence of response-response associations in the data of Table 1. The information-transmission statistic indicates mutual influence of the responses that is almost as large as the transmission between the appropriate signal and the response. A satisfactory analysis of the performance must address this effect. However, at this point there are no

further factors in the design that can be used as the basis of a test of unrelatedness. To gain further understanding, we must move from hypotheses that express the disconnection of two classifications to models that describe the nature of their connection.

To start the analysis we return to the list of interactions in the detection process that we introduced above. One can account for the XY association in several ways—as correlated sensory noise, as a bivariate-response interaction, or as some interaction of causes. A clean separation of these effects would be ideal. However, without a method to record the interior steps of Figure 1, an unambiguous separation is impossible. If either correlated noise or bivariate-response interaction are allowed to take fully general forms, then either could explain any observed result. To make a separation, it is necessary to restrict the forms taken by these effects.

A partial separation of the causes of the response-response association can be made if some additional assumptions about response interaction are introduced. Three assumptions are plausible and drive our analysis below.

1. Sensory noise has a simple relationship to the categories. Following the standard assumptions of bivariate signal-detection theory (e.g., a Gaussian noise distribution) this portion of the association looks like linear correlation when related to a scaling of the categories that preserves their order.

2. Effects that involve numerical bias, being more cognitive in nature, are related by some simple rule to the actual numbers used by the observer. They do not have the bilinear form of sensory noise.

3. Strange-looking associative effects are most likely to be the result of an inadequate separation of sensory and numerical effects. There should be no distributions with isodensity contours that look like beans, waffles, or other improbable objects (cf. Klein, 1985).

These three assumptions suggest how the association effects might be separated. First, fit a model to the data that allows for direct stimulus effects and that represents the XY association in bilinear form. Several forms for this model are discussed below. Second, examine the residuals from this model and attempt to identify patterns that are closely related to the use of numbers. Third, exclude these numerical patterns from the data or describe them with a model, then refit the bilinear association models to obtain more accurate estimates of its parameters.

In good part, this program of analysis is based on ambiguous criteria. It is exploratory, not confirmatory, and often needs to be modified. For example, in data that show strong numerical regularities, it may be preferable to look for the patterns of numerical use before fitting the association model. Eventually, any conclusions should be replicable with other data. In particular, we would expect that effects related to the use of numbers should be present in other data from the same observer and that effects related to correlated noise should show consistency over observers.

To examine the response relationships, the data are fit with a series of models that describe specific forms of association. Thus, we will next describe a class of models appropriate to the association patterns of bivariate rating data. We will then return to the analysis of Table 1.

The Multiplicative-Association Model

There are many ways to represent an ordinal pattern of XY association. Researchers familiar with signal-detection models in psychophysics will find the bivariate Gaussian distribution to be the most familiar example of such a model. This model imposes a bivariate distribution on an underlying continuum in which the joint percept is represented. The general bivariate normal distribution has five parameters: two means (μ_x and μ_y), two variances (σ_x^2 and σ_y^2), and a correlation (ρ). The correlation measures the association between the responses. Using this model, one can approach Table 1 by fitting the bivariate normal distributions to each of the HL combinations, estimating the parameters, then comparing the values of ρ across the various stimulus conditions.

This description is natural and has the advantage that it is relatively easy to interpret. However, it suffers from three important defects relating to its computational aspects. First, the parameters of a categorized bivariate normal distribution are not easy to estimate. The estimation problem has been solved—the correlation here is conventionally known as the *polychoric correlation*—but the associated algorithms are not simple (Beardwood, 1977; Martinson & Hamdan, 1975). A correlation coefficient based on the numerical category ratings—the integers 1 through 6 here—cannot be calculated simply since there is no assurance that the responses are equally spaced with respect to the Gaussian distribution.³ Second, it is difficult to extend this model, particularly its parametric-estimation scheme, to multiple stimuli. This difficulty makes it hard to fit models and examine hypotheses that describe several stimulus conditions at once. Finally, it is hard to accommodate exceptional cells into the analysis or to conjoin a Gaussian-association model with a model for numerical bias. For these reasons, and to maintain continuity with the association-based analysis, we turn to a different class of description for these data.

An effective choice for the present problem is a model known as the *multiplicative-association model* (Goodman, 1979). An introduction to this model is provided in Appendix A-5. To summarize that explanation, for a two-way contingency table the multiplicative-association model depends on six types of parameters: (1) a parameter κ that depends most directly on the overall sample size, (2) a set of parameters α_x associated with the relative frequency of observations in the rows of the table, (3) a comparable set of parameters β_y associated with the column frequency, (4) a set of parameters ξ_x that express the scaling of the rows with respect to the association, (5) a comparable set of parameters η_y for the columns, and (6) a single parameter ϕ that indicates the magnitude of the bilinear association of the ξ_x to the η_y . As a description of the expected frequencies in the table, the model takes the form of a

log-frequency equation in which a product term is appended to the usual additive relationship under independence,

$$\log \mu_{xy} = \kappa + \alpha_x + \beta_y + \phi \xi_x \eta_y.$$

The association component of this model is product term $\phi \xi_x \eta_y$. This form, which is linear in both ξ_x and η_y , closely approximates the association in the Gaussian distributions postulated by signal-detection theory (see Appendix A-5b). The relationship between the correlation ρ in the bivariate Gaussian model and the association parameter ϕ of the multiplicative-association model is

$$\phi \approx \frac{\rho}{1 - \rho^2} \text{ or } \rho \approx \frac{\sqrt{1 + 4\phi^2} - 1}{2\phi}$$

(see Equations 24 and 25). Unlike the direct Gaussian model, the multiplicative-association model is readily applied to frequency-table data.

The parallel to the Gaussian model allows the multiplicative-association model to be used to represent the detection data. When so applied, the scale values ξ_x and η_y roughly indicate the centers between the cutpoints that define the categories. The multiplicative-association model is more general than a categorized bivariate Gaussian distribution, since it does not impose a relationship between the marginal-frequency parameters, α_x and β_y , and the association-scale parameters, ξ_x and η_y . In the true Gaussian model, a single set of cutpoints determines both marginal distributions and the association parameters (see Equations 21–23). We have found that free estimations of ξ_x and η_y often lead to instability of the parameter estimates. Accordingly, in our analysis we assign values to these parameters a priori (see Appendix A-5a). We consider two versions of the model. In the uniform-spacing model the parameters are chosen to reflect equal association between adjacent categories. In the logistic-spacing model they are chosen to reflect the width of the categories, by extracting them from a Gaussian distribution fit to the marginal distributions.

As described, the multiplicative-association model applies to the two-dimensional table from a single stimulus condition. However, the concurrent detection ratings in Table 1 also depend on the four stimulus conditions. For our analysis of this task, we wish to fit the model simultaneously to the full table. In doing so, we are free to let the parameters of the model vary across the stimuli or to hold them to a single value. The parameters of the marginal distributions, κ , α_x , and β_y , are not of central concern, so we allow them to vary with the stimulus condition as needed to fit the data best. The values of the scales, ξ_x and η_y , are fixed by the spacing model mentioned above and do not vary. Appending the subscripts h and l to indicate signal variation, our model is

$$\log \mu_{hly} = \kappa_{hl} + \alpha_{hix} + \beta_{hly} + \phi_{hl} \xi_x \eta_y.$$

Variation of the parameter ϕ , which measures the magnitude of bilinear association is of substantial interest. To

examine it, we consider two versions of the multiplicative-association model. In the *heterogeneous association* version we allow ϕ to take different values in each stimulus condition, and in the *homogeneous association* version we constrain it to a single value. Thus, in all we investigate four association models, depending on (1) whether ϕ is fixed or free, and (2) whether ξ_x and η_y have a uniform or a logistic spacing.

The Multiplicative-Association Model for Concurrent Detection

In the concurrent detection data of Table 1, a substantial XY association was deemed to be present. We noted in our analysis of the task that an important portion of this association (the portion due to correlated noise in the channels) should reside in a bilinear component. Thus, we expect that some form of the multiplicative-association model may be a good approximation of the data.

Fits of the multiplicative-association models to the data in Table 1, as expressed by values of G^2 , are shown in Figure 2.⁴ This figure is constructed in the form of a hierarchy diagram, with models at the bottom containing those connected to them from above as special cases. When two models are connected by a line or a series of descending lines, the lower model contains all the parameters of the upper model, along with some extra parameters. The statistical significance of these added parameters is measured by the difference in the G^2 statistics of the two models. As noted in Appendix A-2, these tests are more reliable than tests of overall fit, and are less affected by small cell frequencies.

Three models other than the multiplicative-association model are shown in Figure 2. At the top is the model $X \parallel Y | HL$, the deviations from which are to be explained. In the middle is the model $[HLX][HLY][XY]$, which implies that the pattern of association is the same in each stimulus condition. Unlike the homogeneous- ϕ model, it

does not require the form of the association to be multiplicative. This model would be rejected if the form of the XY association changed appreciably with the different stimuli. At the bottom of the figure is a model labeled "Saturated." This model contains as many parameters as there are data cells. It fits perfectly and tests nothing; it is the endpoint that ties the hierarchy of models together.

Several conclusions follow from the results in Figure 2. First, none of the multiplicative-association models fits very well. This finding suggests that models based on a single latent noise distribution are unsatisfactory. Second, there is still some evidence for correlation effects, as indicated by a change of $G^2 = 13.8$ between $X \parallel Y | HL$ and the uniform-spacing model, and a comparable change of $G^2 = 12.0$ for the logistic-spacing model. Third, there remains a substantial association effect, since the difference in G^2 between either the uniform-spacing model or the logistic-spacing model and $[HLX][HLY][XY]$ is substantial.

The situation with respect to associational homogeneity over the stimuli is less clear. As noted earlier, the model $[HLX][HLY][XY]$ fits reasonably well, the G^2 statistic not being sufficient to reject it. This indicates that one need not look for heterogeneity of association. However, the differences between the homogeneous and the heterogeneous versions of either the uniform-spacing model or the logistic-spacing model are significant ($G^2_3 = 37.0$ and 43.4 , respectively). The sample estimates confirm this heterogeneity. Estimates of ρ_{ni} for the uniform-spacing model, are 0.12, 0.04, -0.24, and -0.28, for conditions *aa*, *ap*, *pa*, and *pp*, respectively, including a correction for categorization (Equation 26). Together, these results suggest that homogeneous association is tenable only if one abandons the assumption that the association involves only Gaussian noise.

Deviations from the Multiplicative-Association Model

The failure of the simple multiplicative-association models to fit the data in Table 1 must be attributed to causes of association other than bivariate noise. According to our assumptions, one source of these effects may be response patterns. One way to look for these patterns is to examine the residuals from a fit of one of the bilinear association models. The homogeneous version of the uniform-spacing model is a good starting point. It is fitted, expected frequencies are found, and the standardized residuals are calculated (see Appendix A-3, Equation 9). These are examined for sensible numerically-related patterns.

Table 5 shows these deviations. In such a table, it is useful to flag the most deviant cells. As in Table 3, significance tests *per se* are not useful here, but marking roughly 10% of the most deviant cells is valuable. In Table 5, boxes have been drawn around cells with $|d_{hly}| > 3$. The most prominent deviations are in the cells where $X=1$ and $Y=1$. This excess is most striking in the *a-a* condition, but positive effects are also evident in the other tables. The remaining cells do not show such an ob-

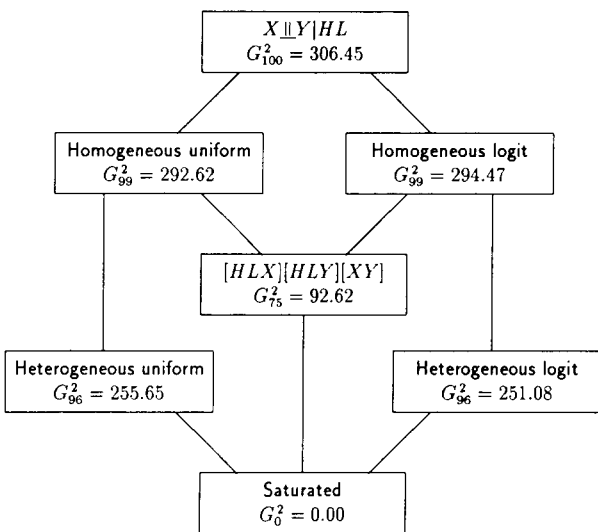


Figure 2. Hierarchy diagram showing the fits of association models to the data from Table 1.

Table 5
Standardized Deviations from the Homogeneous-Association Uniform-Spacing Model Applied to the Data in Table 1

Stimulus Y	X						
	1	2	3	4	5	6	
H = a	1	6.46	-3.12	-1.46	-2.10	-0.87	1.06
	2	-1.97	2.00	0.65	-2.24	1.54	0.94
	3	-2.13	1.70	1.15	1.33	-1.33	-1.79
	4	-0.66	0.20	-0.66	2.58	-1.52	-0.43
	5	-1.50	-1.40	1.33	2.18	1.03	-1.06
	6	-0.44	-0.03	-1.46	-0.59	3.54	1.07
L = p	1	3.28	-0.71	-0.81	-1.38	-1.40	1.47
	2	0.08	-0.30	-0.35	-0.33	1.17	0.09
	3	0.89	1.31	0.20	-0.88	-1.17	-1.07
	4	-0.16	0.91	0.61	0.39	-1.49	-1.29
	5	-1.39	-0.38	0.55	0.89	1.33	-0.66
	6	-0.21	-0.55	-0.50	0.19	0.65	1.40
H = p	1	0.56	-1.51	-1.31	-2.21	-1.70	3.36
	2	-0.86	-1.04	0.74	0.08	2.88	-1.60
	3	0.90	0.58	1.06	0.99	-0.42	-1.49
	4	-0.30	2.44	-0.27	2.17	-1.44	-1.30
	5	0.42	0.85	-0.51	-0.72	0.07	-0.05
	6	-0.83	-0.90	0.36	0.21	-0.11	0.81
L = a	1	1.52	-1.18	-0.85	-2.43	-1.70	2.55
	2	-1.33	0.77	-1.76	-1.75	0.30	1.55
	3	-0.72	0.00	-0.14	0.97	0.41	-0.47
	4	-0.11	-0.81	0.71	2.70	-0.04	-1.40
	5	-1.08	1.65	0.20	0.24	1.82	-1.82
	6	1.26	-0.56	0.88	-0.35	-0.84	0.02

Note—Boxed cells indicate deviations greater than ± 3 .

vious pattern. Negative deviations elsewhere in the $X=1$ column and $Y=1$ row are likely to be a consequence of the fact that the fitting algorithm forces the observed and expected row marginals to agree, so that an excess of observations in one cell is coupled with deficiencies elsewhere. These observations suggest a response pattern in which denials of both signals occur too often.

Having identified a potential source of association, we next check to see how its removal affects the model. An easy way to do this is to exclude the suspect cells from the test. One of the advantages of working with the multiplicative-association model is that cells can be treated as "structural zeros" that are removed from the analysis. The specified cells make no contribution to either the estimation of the parameters or to the test of the models. When the same model is fitted to complete and incomplete tables, the results are hierarchically related, so that differences in G^2 indicate the extent to which the deleted cells are a substantial anomaly.⁵

When the cells with responses $X=Y=1$ are deleted from the analysis and the models for bilinear association are

fitted, the statistics in Figure 3 are obtained. These increase the quality of the analysis in several respects. First, there is a general improvement in the fit, with G^2 dropping substantially for all models except $[HLX][HLY][XY]$. Thus, the four cells are, indeed, a significant anomaly with respect to the single-component association models. Second, the greatest beneficiaries of this improvement are the homogeneous-association models. For example, G^2 for the homogeneous logistic-spacing model drops by 92.2 on 4 degrees of freedom. As measured by the differences in their G^2 statistics, these models are now much further from conditional independence ($G^2_1 = 46.9$ for the logistic-spacing model) and much closer to the heterogeneous-association models ($G^2_3 = 17.4$). Although the latter statistic is still significant, it indicates much less variation in ϕ . When converted to estimates of the bivariate correlation ρ , the four association values are -0.10 , -0.02 , -0.28 , and -0.34 . Under the homogeneous-association model, these are replaced by the common estimate -0.25 . This value is evidence for a modest source of negatively correlated noise. This constant association is not apparent in an analysis of the data at a more superficial level. The negative association can be given a psychophysical interpretation and appears to be of substantive importance (Hirsch, Hylton, & Graham, 1982; Olzak, 1986).

At this point one might begin another cycle of residual examination to try to develop a more elaborate model for response tendencies. Unfortunately, no particularly obvious patterns are present. It appears that whatever association remains in Table 1 is distributed over the complete table. Although models for this pattern of association can be constructed, we do not follow them here.

A more productive approach is to turn to other data to see whether a similar result is obtained. The experiment from which the data in Table 1 were recorded included other conditions that involved different spatial frequency

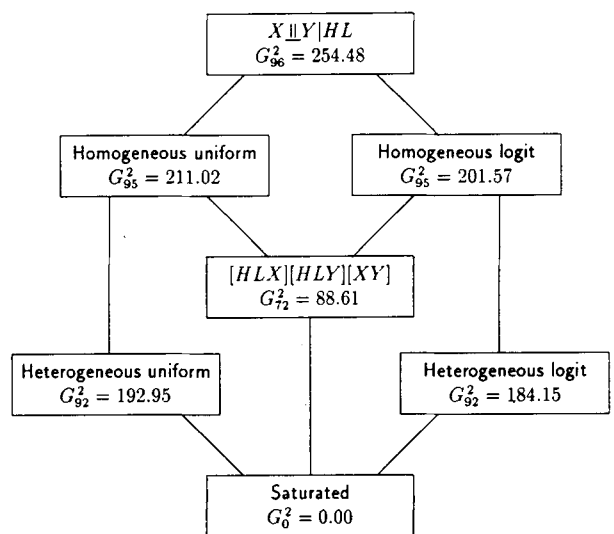


Figure 3. Hierarchical diagram of the fits of association models to the data from Table 1 with cell (1,1) treated as a structural zero.

signals and different exposure times. An examination of these conditions shows a remarkably similar pattern of results. Response-response dependence is clearly present, this pattern appears relatively homogeneous over the stimulus conditions (although not completely so), and the (1,1) response cells are a significant anomaly in the data, the removal of which makes the remaining analysis more consistent.

SUMMARY

Our analysis of Table 1 has identified five influences on the responses:

1. The detection of H , indicated by its association with X .
2. The detection of L , indicated by the comparable LY association.
3. A negative effect of the high-frequency signal on the low-frequency response, indicated by the rejection of the hypothesis $H \perp Y | L$. In terms of our picture in Figure 1, this association is viewed as an inhibitory influence on the input channel. There is no comparable influence of the low-frequency signal on the high-frequency responses.
4. A tendency for the subject to indicate a complete absence of both responses more often than is consistent with the remainder of the responses. We believe this was a response strategy and expect it to appear in other data for this subject.
5. A modest negative interaction of the responses. In the representation of Figure 1, this results from the presence of negatively correlated noise in the combined percept. This noise source does not depend on the stimulus.

This analysis provides a relatively fine-scale partitioning of the interactions in the signal-detection task. We emphasize that it would not have been possible without the use of the concurrent detection task and without the detailed fitting of statistical models. Although we recognize that some aspects of our interpretation may be subject to argument (particularly those related to the analysis in Figure 1), we believe that the picture we have obtained is both more simple and more clear than a less elaborate analysis could have provided.

REFERENCES

- AGRESTI, A. (1984). *Analysis of ordinal categorical data*. New York: Wiley.
- AGRESTI, A., & YANG, M.-C. (1987). An empirical investigation of some effects of sparseness in contingency tables. *Computational Statistics & Data Analysis*, *5*, 9-21.
- ASHBY, F. G. (1988). Estimating the parameters of multidimensional signal detection theory from simultaneous ratings on separate stimulus components. *Perception & Psychophysics*, *44*, 195-204.
- ASHBY, F. G., & TOWNSEND, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154-179.
- BEARDWOOD, J. E. (1977). A remark on algorithm AS 87: Calculation of the polychoric estimate of correlation in contingency tables. *Applied Statistics*, *26*, 121.
- GARNER, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: Wiley.
- GOODMAN, L. A. (1979). Simple models for the analysis of association in cross-classification having ordered categories. *Journal of the American Statistical Association*, *74*, 537-552. [Reprinted in Goodman, 1984]
- GOODMAN, L. A. (1981). Association models and the bivariate normal for contingency tables with ordered categories. *Biometrika*, *68*, 347-355. [Reprinted in Goodman, 1984]
- GOODMAN, L. A. (1984). *The analysis of cross-classified data having ordered categories*. Cambridge, MA: Harvard University Press.
- GOODMAN, L. A. (1985). The analysis of cross classified data having ordered and/or unordered categories: Association models, correlation models, and asymmetry models for contingency tables with or without missing entries (1983 Henry L. Reitz Memorial Lecture). *Annals of Statistics*, *13*, 10-69.
- GREEN, D. M. (1983). Profile analysis. *American Psychologist*, *38*, 133-142.
- GREEN, D. M., MASON, C. R., & KIDD, G., JR. (1984). Profile analysis: Critical bands and duration. *Journal of the Acoustical Society of America*, *75*, 1163-1167.
- HABERMAN, S. J. (1977). Log-linear models and frequency tables with small expected cell counts. *Annals of Statistics*, *5*, 1148-1169.
- HABERMAN, S. J. (1981). Tests for independence in two-way contingency tables based on canonical correlation and on linear-by-linear interaction. *Annals of Statistics*, *9*, 1178-1186.
- HIRSCH, J., HYLTON, R., & GRAHAM, N. (1982). Simultaneous recognition of two spatial frequency components. *Vision Research*, *22*, 365-375.
- HOWARD, E., MACKINNON, D. P., WOODWARD, J. A., & THOMAS, J. P. (1984). Single subject ANOVAS: Testing and correcting for dependent observations. *Investigative Ophthalmology and Visual Science*, *25*(Suppl. 3), 295. (Abstract)
- KENDALL, M. G., & STUART, A. (1977). *The advanced theory of statistics: Vol. 1. Distribution theory* (4th ed.). London: Griffin.
- KLEIN, S. (1985). Double judgment psychophysics: Problems and solutions. *Journal of the Optical Society of America A*, *2*, 1560-1585.
- MARTINSON, E. O., & HAMDAN, M. A. (1975). Calculation of the polychoric estimate of correlation in contingency tables. *Applied Statistics*, *24*, 272-278.
- MCGILL, W. J. (1954). Multivariate information transmission. *Psychometrika*, *19*, 97-116.
- OLZAK, L. A. (1985). Interactions between spatially tuned mechanisms: Converging evidence. *Journal of the Optical Society of America A*, *2*, 1551-1559.
- OLZAK, L. A. (1986). Widely separated spatial frequencies: Mechanism interactions. *Vision Research*, *26*, 1143-1153.
- OLZAK, L. A., & KRAMER, P. (1984). Inhibition between spatially tuned mechanisms: Temporal influences. *Journal of the Optical Society of America A*, *1*, 1290. (Abstract)
- OLZAK, L. A., & THOMAS, J. P. (1986). Seeing spatial patterns. In K. R. Boff, L. Kauffman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1). New York: Wiley.
- OLZAK, L. A., & WICKENS, T. D. (1983). The interpretation of detection data through direct multivariate frequency analysis. *Psychological Bulletin*, *93*, 574-585.
- OLZAK, L. A., WICKENS, T. D., & GOULED-SMITH, B. (1985). Ruling out response bias in the concurrent detection task. *Investigative Ophthalmology & Visual Science*, *26*(Suppl. 3), 142. (Abstract)
- SCHARF, B., & BUUS, S. (1986). Audition I: Stimulus, physiology, thresholds. In K. R. Boff, L. Kauffman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1). New York: Wiley.
- SHERRICK, C. E., & CHOLEWIAK, R. W. (1986). Cutaneous sensitivity. In K. R. Boff, L. Kauffman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1). New York: Wiley.
- THOMAS, J. P. (1985). Detection and identification: How are they related? *Journal of the Optical Society of America A*, *2*, 1457-1467.
- WICKENS, T. D. (1989). *Multiway contingency table analysis for the social sciences*. Hillsdale, NJ: Erlbaum.

NOTES

1. These are not the same data that were previously analyzed by Olzak and Wickens (1983).

2. The reader should not confuse the numerical frequency with which a particular response was made with the spatial frequency of a stimulus component.

3. Even if appropriate, this estimate is biased downward due to the effects of categorization (Ashby, 1988). Some recovery from this bias is obtained from Sheppard's correction, which is discussed with respect to other estimates below (Equation 26).

4. The calculations in this paper were done on an IBM PC. Source and executable versions of the programs, written in TurboPascal (Borland International, Inc.), can be obtained by forwarding a diskette to the first author.

5. An excess of frequency in a selected set of cells can also be accommodated by incrementing all the cells by the same amount. This is a simpler model than that created by excluding the cells, since it differs from the unadjusted model by only a single degree of freedom. However, the psychological interpretation of the multiplicative increment is unclear, and we will not pursue it.

6. A more precise notation for the marginal sums replaces the summed subscripts by dots or pluses, writing the marginal distribution for $U=u$ as f_{u+} , and the marginal distribution for $V=v$ as f_{+v} . The informal notation of Equation 1, which drops the summed subscripts, is less formidable and suffices for this paper. Another convenience of the present notation is that the formulas do not depend on the number of extraneous factors; thus, Equation 1 gives expected frequencies for the hypothesis $U \perp V$ in a k -way table for any k .

7. Some evidence for this contention has been collected by Emily Howard at UCLA, who examined a sequence of identification observations for sequential correlations. She found that the distributional assumptions were satisfied for practiced subjects but failed for naive subjects (Howard, MacKinnon, Woodward, & Thomas, 1984).

8. See Agristi & Yang (1987) for an illustration of this effect for the multiplicative-association model.

9. In these papers, Goodman refers to the multiplicative-association model as the row-and-column effect Model II.

10. Of course, not all pairs are necessary to determine the association; for example, the $(I-1)(J-1)$ tables formed from adjacent pairs of rows and adjacent pairs of columns are sufficient.

11. Goodman refers to this model as the uniform-association model, since the association between adjacent categories is always the same. Our designation emphasizes the scaling of the association.

12. Adjustment equations for the multiplicative-association model when both ξ , and η , are free are given by Goodman (1979). The resulting model is not log-linear, and must be analyzed somewhat differently. Specifically, the reference of differences in X^2 or G^2 to a chi-square distribution under the hypothesis of no association is no longer correct; a Wishart root distribution is required (Haberman, 1981).

APPENDIX

Statistical Theory

In this section we describe the statistical basis of our analysis, both for readers for whom some of our analysis is unfamiliar and for those who wish to duplicate the analysis. For a more extensive discussion of the statistical models, see Agristi (1984) or Wickens (1989).

A-1. Independence and Conditional Independence

A fundamental hypothesis in a two-way contingency table is the unrelatedness of the row and column categorizations. If both attributes have been sampled, this lack of association corresponds to probabilistic independence, whereas if the row classification is determined as part of an experimental design, unrelatedness represents homogeneity of the columns. We will not be partic-

ular about this distinction here and will use the term *independence* to refer to both situations. For a table with rows indicated by the factor U and columns indicated by the factor V , we denote this lack of association by $U \perp V$.

Under the null hypothesis $U \perp V$, the expected cell frequencies conditional on the observed marginal distributions of U and V are calculated by the usual formula for a chi-square test

$$\hat{\mu}_{uv} = \frac{f_u f_v}{N}. \quad (1)$$

Here f_u and f_v are the marginal frequencies for categories $U=u$ and $V=v$, N is the total number of observations, and $\hat{\mu}_{uv}$ is the expected frequency in the uv th cell.⁶ These expected frequencies are compared to their observed counterparts using statistics such as the Pearson goodness of fit statistic,

$$X^2 = \sum_{\text{cells}} \frac{(f_i - \hat{\mu}_i)^2}{\hat{\mu}_i}, \quad (2)$$

or the slightly less familiar likelihood ratio statistic,

$$G^2 = 2 \sum_{\text{cells}} f_i \log \frac{f_i}{\hat{\mu}_i}. \quad (3)$$

The generic index i in these formulas runs over all cells in the table. As test statistics, either X^2 or G^2 are satisfactory, although because of its relationship to information theory (see Appendix A-4) some interpretative advantage attaches to G^2 . To simplify our exposition, we sometimes attach a hypothesis to its test statistic, for example, by writing $G^2 (U \perp V)$.

To complete these tests as conventionally run, the statistics are compared to a chi-square distribution with degrees of freedom equal to the product of the number of rows minus 1 multiplied by the number of columns minus 1. We denote the degrees of freedom by subscripts, for example by writing X^2_{\dagger} and G^2_{\dagger} . The use of the chi-square reference distribution requires certain assumptions to be made about the data. First, the observations must be independent and identically distributed. For psychophysical data of this sort, these assumptions are not automatically satisfied. The data we considered were collected from a single subject during several sessions, thus raising questions about the independence of observations and their homogeneity from day to day. Although it is difficult to assess the magnitude and effect of these violations, we believe that the assumptions were moderately well satisfied. Trial-to-trial independence is a standard assumption for psychophysical data, and the inhomogeneities, while potentially more serious, are probably minor.⁷ If independence is approximately satisfied, the total frequency in the table must also be large before a chi-square approximation can be used. In general, it is desirable for most of the cells to have expected frequencies appreciably larger than 1. Our data did not always satisfy this requirement. Accordingly, in this paper we calculate the test statistics, but approach their interpretation with caution, particularly as tests of the overall fit of a model. We use the statistical values primarily as indications of relative fit and do not consider effects to be real simply because they are statistically significant, but require in addition that they be of appreciable size.

When one moves to a multiway table, hypotheses of unrelatedness assume a more complex form. One important hypothesis asserts that two factors, say U and V , are unrelated at each level of another factor, W . The hypothesis $U \perp V$ is simultaneously tested at each level w of W . To indicate the conditional status

Table 6
Expected Frequency Formulas and Degrees of Freedom for the Tests of the Independence and Conditional-Independence Hypotheses in Table 4

Hypothesis or Model	Table	Estimated Expected Frequencies	Degrees of Freedom
$[HL][X][Y]$	$HLXY$	$\hat{\mu}_{hxyz} = f_{hi} f_x f_y / N^2$	$(HL-1)(X-1)(Y-1)$
$H \perp\!\!\!\perp X$	HX	$\hat{\mu}_{hx} = f_h f_x / N$	$(H-1)(X-1)$
$L \perp\!\!\!\perp Y$	LY	$\hat{\mu}_{ly} = f_l f_y / N$	$(L-1)(Y-1)$
$[HL][HX][LY]$	$HLXY$	$\hat{\mu}_{hxyz} = f_{hi} f_{hx} f_{ly} / f_h f_l$	$(HX-1)(LY-1) - (H-1)(L-1)$
$H \perp\!\!\!\perp Y L$	HLY	$\hat{\mu}_{hly} = f_{hi} f_{ly} / f_l$	$L(H-1)(Y-1)$
$L \perp\!\!\!\perp X H$	HLX	$\hat{\mu}_{hlx} = f_{hi} f_{hx} / f_h$	$H(L-1)(X-1)$
$X \perp\!\!\!\perp Y HL$	$HLXY$	$\hat{\mu}_{hxyz} = f_{hix} f_{hly} / f_{hi}$	$HL(X-1)(Y-1)$

of this hypothesis, it is denoted $U \perp\!\!\!\perp V | W$. It is often referred to as a hypothesis of *conditional independence*. A conditional-independence hypothesis is not the same as pure independence: not infrequently data is found for which $U \perp\!\!\!\perp V | W$, but for which $U \perp\!\!\!\perp V$ fails when the table is collapsed over W .

Like the independence hypotheses, the hypotheses of conditional independence determine sets of expected frequencies and, subject to the same concerns about the assumptions, are tested by X^2 or G^2 statistics. The expected-frequency formula is the same as Equation 1, except for a conditioning of all parts on the level of W :

$$\hat{\mu}_{uvw} = \frac{f_{uw} f_{vw}}{f_w} \tag{4}$$

When tested, this hypothesis has $W(U-1)(V-1)$ degrees of freedom.

For reference, Table 6 shows the expected frequencies and degrees of freedom needed for the tests of the conditional-independence models we use in our analysis of association.

A-2. Log-Linear Models

The independence model $U \perp\!\!\!\perp V$ is equivalent to a log-linear model for the expected frequencies. As the product form of Equation 1 implies, this model expresses the expected cell frequencies (μ_{uv}) as the sum of terms associated with the levels of U and V ,

$$\log \mu_{uv} = \lambda + \lambda_{U(u)} + \lambda_{V(v)}. \tag{5}$$

In this equation, all parameters on the right side are denoted by λ , with uppercase subscripts to identify the terms that refer to the U and the V dimensions. The lowercase subscripts, u and v , identify the particular level.

More parameters are present in Equation 5 than are necessary. Without additional constraints, the value of λ is confounded with the average values of $\lambda_{U(u)}$ and $\lambda_{V(v)}$. To separate the parameters, the subscripted parameters are chosen to sum to 0 over each index:

$$\sum_u \lambda_{U(u)} = 0 \text{ and } \sum_v \lambda_{V(v)} = 0. \tag{6}$$

Similar constraints apply to other log-linear models.

The conditional independence hypothesis $U \perp\!\!\!\perp V | W$ is equivalent to the log-linear model

$$\log \mu_{uvw} = \lambda + \lambda_{U(u)} + \lambda_{V(v)} + \lambda_{W(w)} + \lambda_{UV(uv)} + \lambda_{VW(vw)}. \tag{7}$$

The parameters $\lambda_{UV(uv)}$ and $\lambda_{VW(vw)}$ in this model allow association between these factors; the lack of a term that depends simultaneously on U and V implies their conditional independence.

There are many log-linear models that do not describe simple hypotheses of independence or conditional independence,

so a shorthand method to express them is helpful. The most useful such models are hierarchical, in that the presence of a multi-factor term implies the inclusion of all the simpler terms in the same factors. Such models are denoted by listing only their high-order terms. Therefore, the model of Equation 5 is denoted $[U][V]$, while that of Equation 7 is denoted $[UW][VW]$. Not all hierarchical log-linear models have simple formulas for the expected frequencies. However, those used in this paper are simple and are shown in Table 6.

A comparison of the fit of two related models is often valuable. Two models, M_1 and M_2 , are hierarchically related when the parameters of M_1 are a proper subset of the parameters of M_2 . A likelihood-ratio test comparing two such models is obtained by taking the difference in the G^2 statistics that test the quality of their fits,

$$G^2(M_2 | M_1) = G^2(M_1) - G^2(M_2) \tag{8}$$

The number of degrees of freedom for this test is the difference in the degrees of freedom of the tests of the two models. In effect, Equation 8 provides a test of the hypothesis that the parameters of M_2 take the particular values that reduce the model to M_1 .

The difference statistic of Equation 8 obeys different rules for convergence to a chi-square distribution than do the simple G^2 statistics that assess overall fit (Haberman, 1977). The quality of the chi-square approximation depends on the total sample size per degree of freedom of the difference statistic, not on the expected frequencies in the individual cells. Thus, when the sample is large, the test is relatively unaffected by small expected frequencies.⁸ In the present analysis, greater trust can be placed in these tests than on the tests of overall goodness of fit.

A-3. Standardized Residuals

When a model (such as the one implied by a test of unrelateness) is rejected, one has evidence for the existence of an association but no information about its form. One good way to explore the association's form is to examine the way the observed frequencies differ from those expected under the model. This comparison can be made by examining the *standardized residuals*. These quantities measure the error of the prediction relative to the square root of the expected values,

$$d_i = \frac{f_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i}} \tag{9}$$

These deviations are closely related to the contributions of each cell to the test statistics; the Pearson statistic X^2 in Equation 2 is the sum of the d_i^2 .

In order to interpret the standardized deviates, it is helpful to know their sampling distribution. Unless N is small or $\hat{\mu}_i$ is near either 0 or N , this distribution is close to the unit normal,

somewhat less variable if many parameters have been estimated in fitting the model. On this basis, there is a temptation to refer them to standard criteria for significance. Indeed, large deviations, say greater than ± 3 , are likely to be real, but it is unwise to assign too much importance to individual effects. Beyond the usual problems of sample size and the independence and homogeneity of the observations, there are several more difficulties. One problem is simultaneous testing; there is one residual for each cell of the table, so the chance of an accidental Type I error somewhere in the table is much greater than the nominal significance level. Moreover, the residuals are linked together by the model that has been fit, so they are not independent of each other. A substantial deviation in one cell influences the fit in the remaining cells. It is usually more revealing to look at the pattern of deviations over the table than to examine single cells.

A-4. Information Theory and Association

The hypotheses of unrelatedness are closely related to the analysis of the association in terms of information theory (Garner, 1962; McGill, 1954). In this analysis, the information or entropy associated with a distribution V is estimated by

$$H(V) = -\sum_v \hat{P}(V=v) \log_2 \hat{P}(V=v) = -\sum_v \frac{f_v}{N} \log_2 \frac{f_v}{N}, \quad (10)$$

where $\hat{P}(\cdot)$ is the estimated probability of an event. The unit of information is the binary digit, or *bit*. For a classification into C categories, the maximum entropy is obtained when $f_v = N/C$ for all v , in which case $H(V) = \log_2 C$. For the six-level response measures in this study, the maximum is 2.585 bits.

The entropy of the joint distribution of two variables, $H(UV)$, is defined by a formula similar to Equation 10 using the joint proportions, f_{uv}/N , and summing over both variables. For any distribution, $H(UV) \leq H(U) + H(V)$, with equality holding only when U and V are unassociated. The information transmitted from U to V is the difference between the maximum possible joint entropy and the actual value,

$$T(U,V) = H(U) + H(V) - H(UV). \quad (11)$$

If there is little relationship between the two classifications in the U - V table, then $T(U,V) \approx 0$, whereas if the classifications are strongly associated, then $T(U,V)$ nears its maximum value. No more information can be transmitted than is available in the smaller classification, so $T(U,V) \leq \min\{H(U), H(V)\}$. In the present study, transmission between the dichotomous signals and the response is limited to 1 bit.

Some algebraic manipulation of Equations 10 and 11 shows that they are closely related to the likelihood-ratio test statistic under independence, $G^2(U \perp V)$ (Equations 1 and 3), with

$$T(U,V) = \frac{\log_2 e}{2N} G^2(U \perp V) = 0.72135 \frac{G^2(U \perp V)}{N}. \quad (12)$$

Thus, $T(U,V) = 0$ if and only if $U \perp V$, and tests of both hypotheses are provided by the G^2 statistic. Unrelatedness and information transmission are complementary ideas, with unrelatedness indicating a lack of transmission and the information transmitted measuring the magnitude of an association.

In a two-way classification, the *conditional entropy* of one classification given another is the weighted average of the simple entropies at each level of the conditioning factor. For example, if $H(U|v)$ is the entropy of U calculated according to

Equation 10 using only the data at level v of V , then the conditional entropy of U given V is

$$H_v(U) = \sum_v H(U|v) \hat{P}(V=v).$$

The *conditional information transmission*, between classifications U and W given a third factor V is defined by conditioning the terms of Equation 11 on V , to give

$$T_v(U,W) = H_v(UW) - H_v(U) - H_v(W). \quad (13)$$

The conditional information transmission is related to the hypothesis of conditional independence as in Equation 12, $T_v(U,W) = (\log_2 e / 2N) G^2(U \perp W | V)$.

A-5. The Multiplicative-Association Model

To go beyond the tests of unrelatedness, the specific form of an association must be described. The multiplicative-association model is one such description. This model is one of several association models that were proposed by Goodman (1979; see also Goodman, 1981, 1984, 1985).⁹ In these models, a numerical score ξ_u is assigned to each row category and a score η_v to each column category of a two-way table, and the model is formed by adding a multiplicative term to the log-linear model for independence (Equation 5),

$$\log \mu_{uv} = \alpha + \alpha_u + \beta_v + \phi \xi_u \eta_v. \quad (14)$$

The multiplicative-association model in Equation 14 is overparameterized. As with the independence models, the parameters α_u and β_v are identified by restricting their sums,

$$\sum_u \alpha_u = \sum_v \beta_v = 0 \quad (15)$$

(cf. Equation 6). The parameters ξ_u and η_v are not determined as to center or scale (as is evident from the form of Equation 18, below). To identify them (and ϕ), two constraints are necessary. For our data, the most practical standardization fixes their mean and variance with respect to their marginal distributions. We denote the marginal probability distributions in the rows and columns by π_u and π_v , respectively. Then fix the means of ξ_u and η_v to 0,

$$\sum_u \pi_u \xi_u = \sum_v \pi_v \eta_v = 0, \quad (16)$$

and their variances to 1,

$$\sum_u \pi_u \xi_u^2 = \sum_v \pi_v \eta_v^2 = 1, \quad (17)$$

These constraints suffice to identify the parameters uniquely.

To understand the multiplicative-association model, it is helpful to see its implication for the pattern of association as measured by the logarithm of the *odds ratio* or *cross-product ratio*. For a 2×2 table, the odds ratio θ is the ratio of the cells on one diagonal of the table to those on the other. We denote the population cell probabilities by π_{uv} , so that $\mu_{uv} = N\pi_{uv}$; then

$$\log \theta = \log \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}} = \log \frac{\mu_{11}\mu_{22}}{\mu_{12}\mu_{21}}.$$

If there is no association in the table, then the corresponding probabilities in the two rows (or columns) are proportional, so that $\theta = 1$ and $\log \theta = 0$. The statistic is positive when the major

diagonal is disproportionately frequent and negative when the minor diagonal is more heavily weighted.

A table with more than two rows or columns cannot be summarized by a single odds ratio. However, the pattern of association over an $I \times J$ table is completely embodied in the set of 2×2 tables created by choosing all possible pairs of rows and pairs of columns.¹⁰ Under the multiplicative-association model, the log-odds-ratio that involves rows u_1 and u_2 and columns v_1 and v_2 is proportional to the differences between the scale values,

$$\log \theta = \phi(\xi_{u_1} - \xi_{u_2})(\eta_{v_1} - \eta_{v_2}). \quad (18)$$

Thus, a table that involves very similar rows has $\xi_{u_1} \approx \xi_{u_2}$ and $\log \theta \approx 0$. A similar argument applies to the columns. One can think of the ξ_u and η_v as scalings of the rows and columns based on the association and of ϕ as an indicator of the overall magnitude of association.

A-5a. Scale spacing. The quantities ξ_u and η_v in the multiplicative-association model determine the spacing of the categories. Several versions of the model can be constructed, depending on how these quantities are chosen. One can leave them as free parameters, to be estimated from the data, but this is not a very satisfactory solution for the present models. We wanted the order of the states to agree with the order of the categories, which is not assured when the estimates are free. A more sensible approach assigns ξ_u and η_v on a priori grounds. The simplest model spaces the categories evenly by assigning ξ_u and η_v , subject to Equations 16 and 17, so that $\xi_{u+1} - \xi_u$ are the same for all u , as are $\eta_{v+1} - \eta_v$ for all v . We refer to this uniformly spaced multiplicative-association model as the *uniform-spacing model*.¹¹

If the categories are used unequally, then one might expect this characteristic to be reflected in the scale spacing. Accordingly, we consider a second version of the model in which ξ_u and η_v depend on the marginal distributions. The cumulative proportion to the midpoints of the categories, sometimes known as the *ridits*, are

$$r_u = \frac{1}{N} \sum_{k=1}^{u-1} p_k + \frac{p_u}{2}$$

(the sum vanishes when $u = 1$). Category scale values are derived from the ridits by converting them to abscissas of a cumulative distribution. For a distribution with cumulative distribution function $F(x)$, the scale value ξ_u associated with r_u is chosen so that $F(\xi_u) = r_u$. A Gaussian distribution is a logical choice for $F(x)$ here, but such a distribution is computationally difficult. Reference to a logistic distribution, which has a nearly identical cumulative distribution function, is simpler. For this distribution, $\xi_u = r_u/(1-r_u)$. The center and scale of these values are adjusted to agree with the normalization of Equations 16 and 17. A comparable development defines the η_v . We call this model the *logistic-spacing model*. Like the uniform-spacing model, it has one free association parameter, ϕ .

A-5b. The multiplicative-association model and the Gaussian distribution. The multiplicative-association model closely approximates the Gaussian distribution of bivariate signal-detection theory. In both models the association is expressed as a bilinear component on a logarithmic scale, as the following argument demonstrates (see also Goodman, 1981, for a numerical example).

Consider a bivariate distribution whose domain has been divided into rectangular regions by a series of cuts orthogonal

to the axes. By the mean-value theorem, the probability of an observation in one of these rectangular regions is equal to the area of the rectangle multiplied by the value of the density function at an interior point. For a sample of size N , the expected frequency falling into the u, v th rectangle is

$$\mu_{uv} = N(\Delta x_u)(\Delta y_v)f(x_{uv}, y_{uv}), \quad (19)$$

where Δx_u and Δy_v are the dimensions of the rectangle and (x_{uv}, y_{uv}) is the interior point. To a reasonable approximation, x_{uv} and y_{uv} may be replaced by points that depend only on u and v , respectively.

Now insert a bivariate Gaussian distribution into Equation 19. The Gaussian density function is

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_x}{\sigma_x} \right)^2 - 2\rho \left(\frac{x-\mu_x}{\sigma_x} \right) \left(\frac{y-\mu_y}{\sigma_y} \right) + \left(\frac{y-\mu_y}{\sigma_y} \right)^2 \right] \right\}.$$

The argument of the exponential is quadratic in x and y and can be rewritten to separate the term involving the product xy . The exponent becomes

$$\frac{\rho\mu_x\mu_y}{\sigma_x\sigma_y(1-\rho^2)} - \frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_x}{\sigma_x} \right)^2 - \frac{2\rho\mu_y x}{\sigma_x\sigma_y} \right] - \frac{1}{2(1-\rho^2)} \left[\left(\frac{y-\mu_y}{\sigma_y} \right)^2 - \frac{2\rho\mu_x y}{\sigma_x\sigma_y} \right] + \frac{\rho xy}{\sigma_x\sigma_y(1-\rho^2)}.$$

Substitution of these expressions into Equation 19 and taking the logarithm gives

$$\log \mu_{uv} \approx K + A_u + B_v + \frac{\rho x_u y_v}{\sigma_x \sigma_y (1 - \rho^2)}, \quad (20)$$

where

$$K = \log \frac{N}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} - \frac{\rho\mu_x\mu_y}{\sigma_x\sigma_y(1-\rho^2)} + C_A + C_B, \quad (21)$$

$$A_u = \log \Delta x_u - \frac{1}{2(1-\rho^2)} \left[\left(\frac{x_u - \mu_x}{\sigma_x} \right)^2 - \frac{2\rho\mu_y x_u}{\sigma_x\sigma_y} \right] - C_A, \quad (22)$$

and

$$B_v = \log \Delta y_v - \frac{1}{2(1-\rho^2)} \left[\left(\frac{y_v - \mu_y}{\sigma_y} \right)^2 - \frac{2\rho\mu_x y_v}{\sigma_x\sigma_y} \right] - C_B. \quad (23)$$

The arbitrary constants C_A and C_B can be chosen so that A_u and B_v sum to zero, as in Equation 15.

The parallel between the Gaussian model of Equation 20 and the multiplicative-association model of Equation 14 is now apparent. Specifically, ξ_u and η_v are identified with x_u and y_v , and to complete the match of the association term, ϕ is identified with $\rho/[\sigma_x\sigma_y(1-\rho^2)]$. By selecting ξ_u and η_v so that $\sigma_x = \sigma_y = 1$ (the constraints imposed by Equations 16 and 17), then ϕ and ρ completely determine each other, with

$$\phi \approx \frac{\rho}{1 - \rho^2} \quad (24)$$

and

$$\rho \approx \frac{\sqrt{1 + 4\phi^2} - 1}{2\phi}. \tag{25}$$

Goodman (1981) pointed out that because of the categorization of data, the value of ρ calculated from this equation is an underestimate. To correct this bias, he suggested adjusting ρ using Sheppard's correction (see Kendall & Stuart, 1977). For a categorization with U levels of U , V levels of V , and ordered scale values, this correction is

$$\rho_{\text{shep}} = \frac{12(U-1)(V-1)\rho}{\sqrt{[12(U-1)^2 - (\xi_U - \xi_1)^2][12(V-1)^2 - (\eta_V - \eta_1)^2]}}. \tag{26}$$

The relationship of K , A_u , and B_v to α_u , α_v , and β_v is more complicated. We do not attempt to incorporate it into our models. As mentioned in Section A-5a, we use the logistic-spacing model to adapt the spacing of ξ_u and η_v to the marginal distribution.

A-5c. Parameter estimates and hypothesis tests. There are no closed estimates for the parameters of the multiplicative-association model. However, they can be found by a number of iterative schemes. Specifically, we use maximum-likelihood estimates, where the maximization is accomplished by the unidimensional Newton-Raphson iteration (Goodman, 1979; see also Agresti, 1984, Appendix B). Although this procedure does not have the most rapid rate of convergence, it has the substantial advantage of being simple and easy to program.

Consider the general version of the multiplicative-association model in which the marginal frequencies and the association are allowed to vary over stimulus types (index s):

$$\log \mu_{uv} = \alpha_s + \alpha_{su} + \beta_{sv} + \phi_s \xi_u \eta_v. \tag{27}$$

To estimate these parameters, it helps to absorb α_s into α_{su} and β_{sv} , and to rewrite the model in multiplicative form,

$$\mu_{uv} = A_{su} B_{sv} \Phi_s^{\xi_u \eta_v}, \tag{28}$$

where $A_{su} = e^{\alpha_{su}}$, and so forth. The quantities ξ_u and η_v are not free parameters,¹² but can be assigned values based on either the uniform-spacing or the logistic-spacing model, consistent with the constraints of Equations 16 and 17. The free parameters are given neutral initial values, typically $A_u = B_v = \Phi_s = 1$. The Newton-Raphson adjustment of the individual parameters are

$$A_{su}^{\text{new}} = A_{su}^{\text{old}} \frac{f_{su}}{\sum_v \hat{\mu}_{su}}, \tag{29}$$

$$B_{sv}^{\text{new}} = B_{sv}^{\text{old}} \frac{f_{sv}}{\sum_u \hat{\mu}_{sv}}, \tag{30}$$

and

$$\Phi_s^{\text{new}} = \Phi_s^{\text{old}} \frac{1 + \sum_{uv} \xi_u \eta_v (f_{su} - \hat{\mu}_{su})}{\sum_{uv} \xi_u^2 \eta_v^2 \hat{\mu}_{su}}. \tag{31}$$

The estimation algorithm alternates between calculating new values of μ_{uv} using Equation 28 and updating the parameters using Equations 29-31. Each cycle consists of three updates of $\hat{\mu}$ interleaved with adjustments of A , B , and Φ . Once the estimates have stabilized, the parameters are converted to those of Equation 27 by setting $\phi_s = \log \Phi_s$, and using the normalization constraints in Equation 15 to convert $\log A_{su}$ and $\log B_{sv}$ to α_u , α_v , and β_v . For the homogeneous-association model, which has constant ϕ , Equations 29 and 30 apply unchanged, and Equation 31 is replaced by

$$\Phi_s^{\text{new}} = \Phi_s^{\text{old}} \frac{1 + \sum_{uv} \xi_u \eta_v (f_{su} - \sum_s \hat{\mu}_{su})}{\sum_{uv} \xi_u^2 \eta_v^2 \sum_s \hat{\mu}_{su}}. \tag{32}$$

Fits of the multiplicative-association model are tested with the usual X^2 and G^2 statistics. Both the uniform-spacing and the logistic-spacing models differ from the conditional unrelatedness model $U \perp V | S$ by the addition of the parameters ϕ_s . Thus, their degrees of freedom are reduced from those of $U \perp V | S$ by the number of free ϕ_s . The multiplicative-association model reduces to conditional unrelatedness when the ϕ_s vanish. This hierarchy lets one test the hypothesis that $\phi_s = 0$. When the assumptions discussed in Appendix A-2 are satisfied, the difference between $G^2(U \perp V | S)$ and G^2 for the multiplicative-association model has a chi-square distribution with degrees of freedom equal to the number of added parameters.

The same estimation algorithm applies to models from which certain cells have been excluded, as we did in our treatment of the cells with $x=y=1$ in Figure 3. The only change in the procedure is that the deleted cells are excluded from all sums (both explicit and implied) in Equations 29-32.