

## CONTINGENCY JUDGMENT: PRIMACY EFFECTS AND ATTENTION DECREMENT \*

J. Frank YATES and Shawn P. CURLEY

*The University of Michigan, Ann Arbor, USA*

Accepted October 1985

Subjects made judgments concerning the strength and direction of the contingency between two dichotomous variables in a situation in which no contingency actually existed. The judgments exhibited a significant primacy effect. The effects of warning and not warning the subjects that they would be required to recall the frequencies of observed event co-occurrences implied that this primacy effect was due to 'attention decrement' (Anderson 1981). According to this hypothesis, attention to contingency-relevant information diminishes after the subject is exposed to only a small portion of the available information.

Judging the contingency between events, e.g., symptom *S* and disease *D* in medical diagnosis, is an important cognitive activity. The pioneering work of Smedslund (1963) and Jenkins and Ward (1965; Ward and Jenkins 1965) led to the recognition that people exhibit limited proficiency at contingency judgment tasks. Those initial studies, as well as subsequent research, have shed much light on the nature of contingency judgment processes as well as the effectiveness of those processes (see reviews by Beyth-Marom (1982) and Crocker (1981)).

Suppose that events *E* and *F* are those of interest in a given situation. The events are contingent if  $P(E|F) \neq P(E|F')$  or, equivalently,  $P(F|E) \neq P(F|E')$ , where the primes indicate the complementary events. The direction and strength of the contingency is measured by what might be called a 'contingency index',  $\Delta = P(E|F) - P(E|F')$ . Table 1 shows a contingency table whose axes are identi-

\* It is our great pleasure to thank Robert Revnew for his assistance in carrying out the reported research. This work was supported in part by U.S. National Institute of Mental Health Grant MH16892.

Requests for reprints should be sent to J. Frank Yates, 136 Perry Building, Dept. of Psychology, University of Michigan, Ann Arbor, MI 48104, USA.

Table 1  
Standard contingency table cell and frequency labels.

Variable 1	Variable 2		Total
	$F$	$F'$	
$E$	Cell A $a$	Cell B $b$	$a + b$
	Cell C $c$	Cell D $d$	$c + d$
Total	$a + c$	$b + d$	$a + b + c + d$

fied with the variables defined by  $E$  and  $E'$ , on the one hand, and  $F$  and  $F'$ , on the other. The frequencies for the four possible joint events are denoted in the cells by lower-case letters. Assuming that observations are selected at random,  $P(E|F)$  is estimated by the frequency ratio  $a/(a+c)$ , and  $P(E|F')$  by the ratio  $b/(b+d)$ . Thus, a proper contingency judgment should be based upon a comparison of these ratios or, equivalently, the frequency products  $ad$  and  $bc$ .

One reason contingency judgments are suboptimal is that people seem to use improper contingency rules (Arkes and Harkness 1983; Beyth-Marom 1982; Shaklee and Tucker 1980; Smedslund 1963). That is, the criteria by which people judge the relatedness of binary variables do not agree with the statistical convention. Rather than comparing  $a/(a+c)$  to  $b/(b+d)$ , subjects have often appeared to judge contingency on the basis of the frequency  $a$ , the so-called 'Cell A strategy'. In other words, they conclude that events  $E$  and  $F$  are contingent to the extent that they co-occur. Another popular rule involves contrasting frequency  $(a+d)$  to frequency  $(b+c)$ . This 'sum of diagonals strategy' is in essence a comparison of the number of cases consistent with the hypothesized contingency of events  $E$  and  $F$  to the number of cases contradicting that contingency.

Strategy selection and resulting contingency judgments are also affected by how the relevant data are presented (Arkes and Harkness 1983; Shaklee and Mims 1982; Ward and Jenkins 1965). The simplest means by which one might judge the contingency of events  $E$  and  $F$  would be with the aid of a data summary, e.g., a contingency table such as that contained in table 1. However, in a more naturalistic situation, information becomes available case-by-case. Even if a person wished to apply an appropriate contingency judgment rule, doing so with a serial

presentation of cases would impose heavy demands on memory. As expected, contingency judgment is more accurate when based on information presented in summary rather than serial formats. Consistent with the memory demand hypothesis, the Cell A strategy seems to be used more commonly when the subject is confronted with large as compared to small numbers of cases.

Beyth-Marom (1982) has demonstrated another factor which influences the strategy by which relatedness is judged. Contingency judgment is affected by a variable's type, or how the variable is described to the subject. An asymmetric binary variable is such that the name of the variable is closely identified with one of the events constituting the variable, but not the other, e.g., 'present' vs 'absent', when the variable is described as 'Pigmentation'. In contrast, for symmetric variables, the constituent events are of equal status, e.g., 'dark' vs 'light', again when 'Pigmentation' is the relevant variable. Subjects are more likely to use contingency rules relying on only part of the relevant information, e.g., the Cell A strategy, when variables are asymmetric rather than symmetric.

Einhorn and Hogarth (1978) have argued that, in many real-world situations, people often lack some of the information needed to determine contingencies. Suppose, for example, that in table 1 event *E* is 'successful job performance' and event *F* is '*predicted* successful job performance'. Employers almost never hire workers who are expected to perform poorly. So, frequencies *b* and *d* will be unavailable, and a proper contingency assessment will be impossible to achieve.

Even if available, relevant information may have little or no effect on judgments of the relatedness of variables (Alloy and Tabachnik 1984; Jennings et al. 1982; Wright and Murphy 1984). Instead, these judgments are often dictated by people's expectations and intuitive theories about how the variables *ought* to be related. The most dramatic demonstrations of such phenomena are the 'intuitive correlations' that subjects report between clinical symptoms and diagnoses in situations in which no contingencies actually exist (Chapman and Chapman 1967, 1969). Moreover, as Einhorn and Hogarth (1978) have shown, when subjects are able to request any frequency information they desire, they often do not ask for all that is needed to make a legitimate contingency determination. Rather, they seek information that is consistent with, but not decisive for, their initial hypotheses about the potential contingency.

The purpose of the present research was to determine whether contingency judgments are affected by an additional factor beyond those implicated by previous studies. In particular, we tested the proposition that the effect of relevant information on a person's contingency judgments would depend on when that information was presented. We hypothesized that contingency judgments would exhibit a primacy effect, such that information presented early would have greater impact than information made available later. This hypothesis was suggested in part by the personality impression literature. It has been found that impressions of target persons typically are more strongly affected by trait descriptors shown early rather than late in a presentation sequence (Anderson and Barrios 1961; Hendrick and Costantini 1970). In effect, we hypothesized that primacy effects in personality impression formation are only one manifestation of more general judgmental mechanisms.

The reported experiment also tested a proposed explanation for the anticipated primacy effect. Anderson (1981) has reviewed evidence indicating that primacy effects in impression formation are often due to 'attention decrement'. After viewing initial data relevant to whether a target person should be liked or disliked, subjects tend to stop paying attention to the incoming data. Thus, the impression of the target person will necessarily be more strongly influenced by the target's traits that happen to be presented early rather than late. Exactly why attention diminishes over time is unclear. A plausible account for the phenomenon is that human judgment processes operate according to subjective sampling principles. These principles determine how generalizations about populations should be inferred from samples drawn from those populations. Formal statistical sampling practice relies upon various forms of the law of large numbers. Tversky and Kahneman (1971) have shown that subjective judgments seem to follow a 'law of small numbers'. If so, then subjects could be expected to stop paying attention to contingency-relevant information quite rapidly, since such effortful attention will be seen as unnecessary for the formation of reasonable judgments.

Our approach to the issues was similar to that taken by Anderson and Hubert (1963) in their study of primacy effects in impression formation. The strategy required that different groups of subjects be shown sequences of data that, overall, implied that the relevant events were unrelated. However, the order of cases contained in each sequence

was such that cases presented early in the sequence indicated a contingency in one direction; cases presented toward the end of the sequence supported a contingency in the opposite direction. In this way, a primacy effect could be identified. Also, all subjects were required to recall the frequencies of case types they observed. However, only half the subjects were forewarned of the recall test; the others were not. If the attention decrement hypothesis had credence, then the anticipated primacy effect in contingency judgments should have been weaker for the forewarned subjects.

## **Method**

### *Subjects*

There were 129 paid subjects in the study. They were assigned randomly to the four conditions of the experiment described below.

### *Stimuli*

The subjects were asked to consider a hypothetical plant called the 'Rhododipsia'. Rhododipsia were reported to grow in both Region A and Region B. Rhododipsia were also said to come in either of two colors, light or dark. During the experiment, subjects were shown what were purported to be the results of a random sampling of Rhododipsia from Regions A and B. Each observation was displayed on a slide indicating the specimen's region and whether it was light or dark.

### *Procedure*

The subjects' task was to determine whether and to what extent Rhododipsia region and color were related to each other. They were informed that the basis for their judgments would be the results of a random sampling of Rhododipsia specimens. The subjects, who served in non-interacting groups of three to five individuals, were then shown practice slides illustrating all four types of specimens that might be observed. The subjects were requested to pay close attention to the slides they would see, but to not keep written records. At the end of the experiment, the subjects were asked to recall how many Rhododipsia from each region and of each color they had been shown. Half the subjects were forewarned of this recall test, while the remaining subjects were not.

Fig. 1a is a contingency table describing the distribution of Rhododipsia shown to the subjects. As is apparent, over the entire sequence of Rhododipsia slides, there was no contingency between region and color. The complete slide sequence actually consisted of two blocks of slides. Fig. 1b shows the distribution of Rhododipsia types contained in Block 1, which implies a contingency whereby light-colored plants tended

a) Overall Distribution: No Contingency

		REGION	
		A	B
COLOR	Light	7	7
	Dark	7	7

b) Block 1 Distribution: Light-A Contingency

		REGION	
		A	B
COLOR	Light	5	2
	Dark	2	5

c) Block 2 Distribution: Light-B Contingency

		REGION	
		A	B
COLOR	Light	2	5
	Dark	5	2

Fig. 1. Distributions of fictitious *Rhododipsia* plants according to region and color: (a) overall, (b) Block 1, (c) Block 2.

to be relatively more common in Region A. The distribution of *Rhododipsia* types contained in Block 2 is shown in fig. 1c. The contingency evident in Block 2 was exactly opposite in direction and strength to that indicated in Block 1. The presentation sequence of slides in Block 1 was random. The slide sequence in Block 2 was isomorphic to the Block 1 sequence. Using letters corresponding to the cells shown in table 1, the transformation was: A  $\rightarrow$  B, B  $\rightarrow$  A, C  $\rightarrow$  D, D  $\rightarrow$  C. That is, an observation belonging to Cell A was replaced by one belonging to Cell B, one belonging to Cell B was replaced by one belonging to Cell A, and so on. Half the subjects saw the slides in the order Block 1-Block 2; the others were exposed to the slides in the opposite order.

The design of the experiment was thus a  $2 \times 2$  factorial. The first factor was whether or not the subject was forewarned of the impending recall test. The second was defined by the order of slide blocks.

After being shown all the specimen slides, each displayed for 4 sec, subjects reported their judgments in written form. Inferred contingency judgments were derived from responses to two conditional probability questions. The subject was asked: 'Suppose a *Rhododipsia* plant is from Region A. What is the chance that it is Light in color rather

than Dark?' The subject was told that he or she should report 100% 'if you are absolutely sure the plant would be Light', and 0% 'if you are absolutely sure the plant would NOT be Light'. Similar instructions defined the meaning of 50% and other intermediate responses. A parallel question requested the subject's judgment of the chance that a plant would be light if it had been found in Region B.

Subjects were also asked to indicate how strongly related, if at all, they felt that Rhododipsia region and color were. They did so on a 0 to 6 scale, on which 0 meant that 'the two features are completely UNRELATED to each other', while 6 implied that 'they are PERFECTLY related to each other'. When the rating was anything other than 0, the subject also reported the direction of the perceived relationship.

The memory task was the final one performed by the subjects. Each individual completed a contingency table by entering the numbers of plants of each type recalled to have been shown in the slides.

## Results and Discussion

For each subject, the following inferred judged contingency index was derived:  $\Delta' = P'(\text{Light}|\text{Region A}) - P'(\text{Light}|\text{Region B})$ , where  $P'$  indicates a judged rather than an 'objective' probability. Fig. 2 displays the mean value of  $\Delta'$ , as a function of slide presentation order and recall test warning.

An analysis of variance concerning  $\Delta'$  was performed on differences in  $P'(\text{Light}|\text{Region A})$  and  $P'(\text{Light}|\text{Region B})$  after these probability judgments were subjected to appropriate variance-stabilizing arcsine transformations. As hypothesized, when subjects were not forewarned of the recall test, a primacy effect was evident. The mean of  $\Delta'$  was significantly higher when the block of slides containing a Light-A contingency was presented first rather than second ( $F(1, 125) = 14.78$ ,  $p < 0.001$ ). As expected, according to the attention decrement proposition, there was not a significant primacy effect in the judgments of subjects who were forewarned of the recall test ( $F(1, 125) = 1.12$ , ns); the interaction of the warning and order factors was marginally significant ( $F(1, 125) = 3.76$ ,  $p < 0.06$ ).

The mean relatedness ratings of subjects in Warning and No-Warning conditions were 2.19 and 2.42, respectively. Although this difference is consistent with expectations, it is not statistically reliable ( $F < 1$ ). The fact that subjects' explicit relatedness judgments do not agree with the judgments inferred from their conditional probability assessments is compatible with previous results (cf. Beyth-Marom 1982). Those studies imply that, although subjects have the capacity to report proper con-

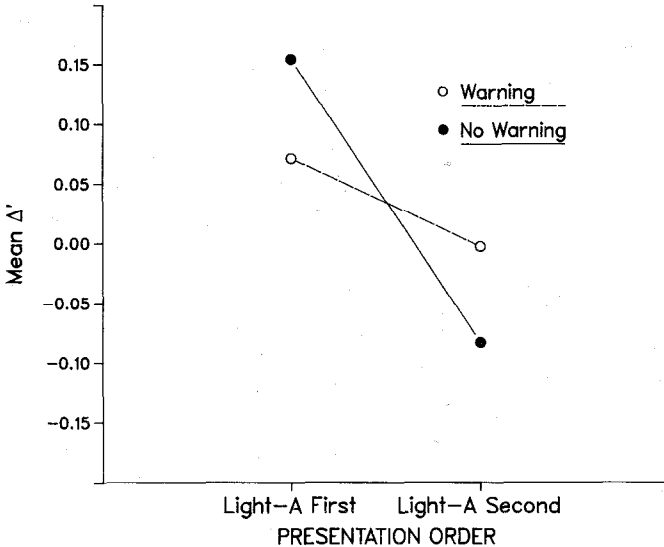


Fig. 2. Mean inferred judged contingency index,  $\Delta' = P'(\text{Light}|\text{Region A}) - P'(\text{Light}|\text{Region B})$ , as a function of presentation order and recall test warning.

tingency judgments, they fail to do so because their subjective notions of relatedness often disagree with the accepted statistical concept.

The accuracy of subjects' recollections of the frequencies of slides showing *Rhododipsia* of various origins and colors were indexed by the following score:

$$S = \sum |f' - 7|,$$

where  $f'$  represents the estimated frequency reported by the subject for each cell in the *Rhododipsia* region-color contingency table, and the summation is over all cells. The mean of  $S$  was 6.31 for subjects in the Warning conditions, and 7.52 for those in the No-Warning conditions. The observed difference is in the expected direction, but not statistically significant. It might also be noted that the subjects were generally quite accurate in estimating the total number of slides they saw. The mean sum of  $f'$  was 28.05 in the Warning conditions, and 27.32 in the No-Warning conditions. These means did not differ significantly from 28, the actual total number of slides, or from each other.



Over the entire sample of specimens shown to each subject, Rhododipsia color and region were completely unrelated. However, if the subject only considered the first half of the sample, that subsample would in fact support a color–region contingency. So a legitimate question to ask is whether, after seeing the first half of the sample, it was appropriate for the subject to have decided that a contingency existed, and therefore stop paying attention to additional sample information. To gain some sense of the reasonableness of such an action, the contingency of Rhododipsia region and color, based on either of the single-block distributions shown in figs. 1b and 1c, was tested according to standard statistical procedures. The resulting (corrected)  $\chi^2$  statistic was 1.14, which is far from conventionally acceptable levels of statistical significance. That is, a contingency judgment based on only the first half of the available information would seem to be a hasty judgment.

In summary, it does indeed appear that contingency judgments are subject to primacy effects. The present data also imply that these primacy effects are due to attention decrement; the ‘law of small numbers’ reaches into yet another quarter. Besides being of interest in themselves, these conclusions offer initial hypotheses for why other puzzling phenomena occur.

Investigators have been struck by the lack of agreement in the judgment policies of different people – including experts – operating in the same domains. For example, Hoffman et al. (1968) studied how nine gastroenterologists made ulcer malignancy judgments on the basis of seven cues discernible from X-rays, e.g., ulcer contour and the presence of related ulcers. These specialists differed markedly in their reliance upon the various indicators. The observed policy differences seemed to be a major cause for the low interjudge agreement; the median correlation between the diagnoses of pairs of physicians was only 0.38. It is conceivable that the types of primacy effects observed in the present study contribute to such policy differences. Natural sampling mechanisms imply that small samples of cases can differ from one another substantially with respect to the contingencies between ulcer characteristics and their malignancies. When neophyte physicians begin learning to diagnose malignancy, the present results suggest that they might be overly (and permanently) influenced by the chance contingencies observed in the first few cases they consider. A plethora of divergent judgment policies would result.

## References

- Alloy, L.B. and N. Tabachnik, 1984. Assessment of covariation by humans and animals: the joint influence of prior expectations and current situational information. *Psychological Review* 91, 112–149.
- Anderson, N.H., 1981. Foundations of information integration theory. New York: Academic Press.
- Anderson, N.H. and A.A. Barrios, 1961. Primacy effects in personality impression formation. *Journal of Abnormal and Social Psychology* 63, 346–350.
- Anderson, N.H. and S. Hubert, 1963. Effects of concomitant verbal recall on order effects in personality impression formation. *Journal of Verbal Learning and Verbal Behavior* 2, 379–391.
- Arkes, H.R. and A.R. Harkness, 1983. Estimates of contingency between two dichotomous variables. *Journal of Experimental Psychology: General* 112, 117–135.
- Beyth-Marom, R., 1982. Perception of correlation reexamined. *Memory and Cognition* 10, 511–519.
- Chapman, L.J. and J.P. Chapman, 1967. Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology* 72, 193–204.
- Chapman, L.J. and J.P. Chapman, 1969. Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology* 74, 271–280.
- Crocker, J., 1981. Judgment of covariation by social perceivers. *Psychological Bulletin* 90, 272–292.
- Einhorn, H.J. and R.M. Hogarth, 1978. Confidence in judgment: persistence of the illusion of validity. *Psychological Review* 85, 395–416.
- Hendrick, C. and A.F. Costantini, 1970. Effects of varying trait inconsistency and response requirements on the primacy effect in impression formation. *Journal of Personality and Social Psychology* 15, 158–164.
- Hoffman, P.J., P. Slovic and L.G. Rorer, 1968. An analysis-of-variance model for the assessment of configural cue utilization in clinical judgment. *Psychological Bulletin* 69, 338–349.
- Jenkins, H.M. and W.C. Ward, 1965. Judgment of contingency between responses and outcomes. *Psychological Monographs: General and Applied* 79 (1, Whole No. 594).
- Jennings, D.L., T.M. Amabile and L. Ross, 1982. 'Informal covariation assessment: data-based versus theory-based judgments'. In: D. Kahneman, P. Slovic and A. Tversky (eds.), *Judgment under uncertainty: heuristics and biases*. New York: Cambridge University Press. pp. 211–230.
- Shaklee, H. and M. Mims, 1982. Sources of error in judging event covariations: effects of memory demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 8, 208–224.
- Shaklee, H. and D. Tucker, 1980. A rule analysis of judgments of covariation between events. *Memory and Cognition* 8, 459–467.
- Smedslund, J., 1963. The concept of correlation in adults. *Scandinavian Journal of Psychology* 4, 165–173.
- Tversky, A. and D. Kahneman, 1971. The belief in the 'law of small numbers'. *Psychological Bulletin* 76, 105–110.
- Ward, W.C. and H.M. Jenkins, 1965. The display of information and the judgment of contingency. *Canadian Journal of Psychology* 19, 231–241.
- Wright, J.C. and G.L. Murphy, 1984. The utility of theories in intuitive statistics: the robustness of theory-based judgments. *Journal of Experimental Psychology: General* 113, 301–322.