

A relative frequency principle in processing contingent information¹

ROBERT G. MILLS, Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson AFB, Ohio 45433

An experiment was conducted to determine the influence on Ss' performance in processing contingent information of: (1) relative frequency (D, the difference in major- and minor-axis cell frequencies of fourfold contingency tables), (2) number of events presented (N, or 20, 30, 40 events), (3) presence or absence of KR information, and (4) several levels of a correlational metric. Six Ss made predictive estimates of the state of a hypothetical system based on observations of sequentially presented categorical information regarding the status of its subsystems. The status information was derived from four-fold contingency tables representing 11 tetrachoric correlation coefficients within the approximate range of ± 1.0 . Results indicate estimates are most accurate as N increases, when based on positive contingencies, and that D can be an appropriate model of Ss' estimate performance.

Erlick & Mills (1967) demonstrated the possibility of a frequency principle in relationship (Pearson r) estimation by applying a transform to Ss' estimates (r_{est}). Their transform, which was based on the linear accumulation of relative weights of variable states, resulted in improved estimate accuracy, and r_{est} functions approached linearity. The implication of their results was that Ss' r_{est} performance was dependent upon the relative frequency of confirming and nonconfirming (with respect to an intuitive hypothesis of degree of relationship) events rather than a least-square principle that would be necessary for accurate r_{est} .

The presence of a relative-frequency principle in r_{est} performance has also been indicated by other investigators (e.g., Inhelder & Piaget, 1958; Smedslund, 1963). However, no investigation to date has provided a convincing test of the utilization of this principle in processing contingent information. The present study attempted to provide such a test by determining if an a priori derived relative-frequency parameter can provide a more adequate description of r_{est}

Fig. 1. Subject display depicting indicator and legend positions.

performance than does a correlational parameter. Also investigated were the influences on r_{est} performance of sample size or number of events presented and the presence or absence of knowledge-of-results information.

METHOD

In this study, Ss were required to make predictive estimates of the state of a hypothetical system (i.e., missile) based on observations of sequentially presented qualitative status information. The simulated environment was depicted as belonging to a missile complex in which maintaining a generalized knowledge of a missile's state of readiness, relative to the operational status of the missile's subsystems, is required. Each sequence-of-event information was presented in terms of subsystem status (two states) as a function of check-out period (also two states) and was derived from fourfold contingency tables containing event frequencies.

Apparatus

The task required Ss to view sequences of binary, go/no-go events presented on a display panel that consisted of four indicator modules arranged horizontally (Fig. 1). One of four possible conditions could be noted with a briefly lighted signal on a given indicator. These conditions were from left to right on the panel: (1) elapsed check-out time greater than 4 h and subsystems status "go," (2) elapsed check-out time greater than 4 h and subsystem status "no-go," (3) elapsed check-out time less than 4 h and subsystem status "go," and (4) elapsed check-out time less than 4 h and subsystem status "no-go."

The display was 12 x 2-3/8 in., with a 2-3/16 in. space between indicator pairs. Each module was 1-1/16 x 7/8 in., with a display surface of 13/16 x 5/8 in. The physical separation between the display surfaces of each pair was 1/8 in. The panel was mounted approximately 18 in. in front of S. Two digital, one-plane modules (3/4 x 5-1/8 in.), for presenting numerical

knowledge of results (KR), were mounted adjacent to each other and directly above the display panel.

Observations were made by S in a three-sided booth, with S's back to the open side. The booth was located in a room separate from all programming and driving equipment.

Stimuli

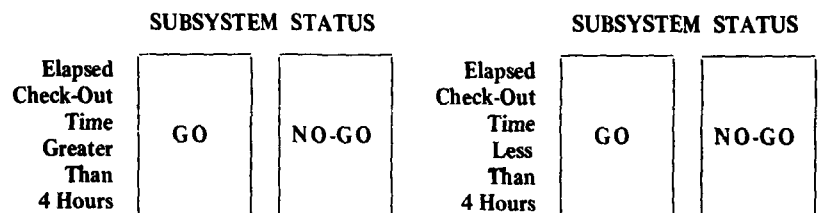
Experimental sequences consisted of a series of events, with each event representing one of the four conditions above. The length of each sequence was defined by the total number of events presented (N). Four-fold contingency tables, containing frequencies of event-condition occurrences and quantified by the tetrachoric correlation coefficient (r_t), were used to determine the sequences. Each r_t table was also defined by one of three N levels (N1 = 20, N2 = 30, N3 = 40 events) and by one of two levels (D1 or D2) of a relative-frequency parameter, $D = (a + d) - (b + c)$, where a, b, c, and d are cell frequencies of the four conditions designated above.

A computer program was used to generate a large quantity of tables varying in D, N, and r_t , from which the experimental tables were selected. The experimental range of r_t presented was ± 1.0 in 0.2 steps or 11 different values. (Since r_t is undefined for ± 1.0 or when any cell frequency equals zero, the experimental values used for $r_t = \pm 1.0$ were actually $r_t = \pm 0.95$.) The values for each D level were selected on the basis of maintaining unique and specific D functions across N and did not in all cases represent absolute minimum or maximum functions. Thus, each sequence presented was defined by a D, N, r_t combination.

A computer program was used to punch a continuous coded paper tape that controlled electronic display equipment. The tape contained coding for 12 experimental sessions of 66 sequences per session or a total of 792 sequences. The order of events within each sequence and the order of the 792 sequences was randomized over the 12 sessions. Each sequence was followed on the tape by a code corresponding to correct r_t value in order that KR could be presented.

Subjects

Six undergraduate university students



made up the S group. All Ss were paid volunteers.

Procedure

The Ss' task was to monitor sequences of qualitative event information conveying subsystem status of a hypothetical missile during check-out. Following each of the sequences, Ss were required to estimate the overall "state of readiness of the missile to fire successfully" (r_{est}). The information monitored was depicted as having been generated by a sequential sampling of *subsystems* that were proceeding at *varying* rates through their 8-h check-out periods. The Ss had no knowledge of what particular subsystem was being sampled at the time each event was displayed.

As a decision aid and for instructive purposes, the following facts concerning past performance of the missile were given.

(1) The degree of readiness for the missile is highest when the status of its subsystems provide either of the following signals: (a) elapsed check-out time is greater than 4 h with a "go" indicated, or (b) elapsed check-out time is less than 4 h with a "no-go" indicated.

(2) The degree of readiness for the missile is lowest when the status of its subsystems provide either of the following signals: (a) elapsed check-out time is greater than 4 h with a "no-go" indicated, or (b) elapsed check-out time is less than 4 h with a "go" indicated.

All Ss were given ample time with instructions and examples in order to insure complete understanding of the task and of the above conditions. At no time were Ss informed of any underlying quantifiable relationship for the information being monitored.

All Ss served under 12 sessions of KR and 12 sessions of no knowledge of results (NKR), or 24 total sessions per S. KR was

displayed on the two modules described above by indicating the sign, + or -, and the degree of readiness, 0 through 10 in integer steps of 2. Three Ss received KR followed by NKR, and the remaining three received NKR followed by KR.

The Ss served individually in daily sessions approximately 45 min in length. Estimates were made in terms of the 11-category scale described for KR display and were recorded on computer transcript sheets following each sequence. Initiation of sequences and recording r_{est} was self-paced. Event information, however, was presented at a rate of 2 events/sec, with an equal on/off ratio.

RESULTS AND CONCLUSIONS

Practice and Order Effects

No systematic changes as a function of practice or order of KR-NKR sessions were found for any of the experimental conditions.

Parameter Effects

All r_t , r_{est} data were analyzed in terms of the KR-NKR, D, and N parameters. Based on this analysis, all parameters except KR-NKR were observed to contribute to r_{est} performance. Figures 2, 3, and 4 are representative r_{est} functions, with D as a parameter in each figure. In these figures, an empirical data point is the mean of six Ss (r_{est}) each making approximately 22 observations for Figs. 2, 3, and 4 and 44 observations for Fig. 5. The use of combined data was based on the fact that individual r_{est} functions were highly representative of r_{est} functions for the group under all experimental conditions. D1 and D2 dashed lines represent the calculated functions (i.e., expected based on the sequences). The solid lines represent r_{est} made to sequences with D1 or D2 values.

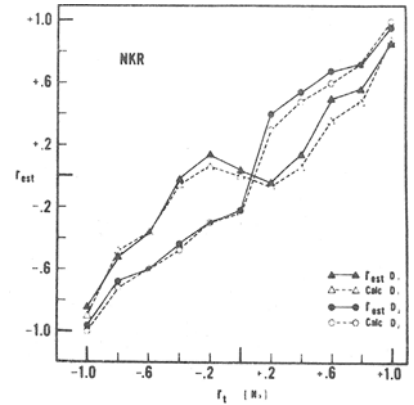


Fig. 4. Partitioned r_{est} functions for sequences of 40 events (N3) under the NKR condition. Functions are partitioned according to two levels of the parameter D.

It is apparent that the empirical r_{est} functions are quite similar to the calculated D functions. It is also apparent upon close examination of these functions that r_{est} is quite sensitive to ordinal or local variations in D that result in functional following of r_{est} across all levels of the experimental parameters. Comparisons of Fig. 4 with the NKR function of Fig. 5 indicate further that considerable disparity results when a r_{est} function is partitioned according to D1 and D2.

An analysis of variance (repeated measures) of mean error for r_{est} indicated the main effect of D to be highly significant ($F = 223.61$, $df = 1/10$, $p < .01$). The main effect of D accounted for 81.8% of the total variance in r_{est} error. There were also main effects of N ($F = 23.21$, $df = 2/20$, $p < .01$) and sequence sign ($F = 21.22$, $df = 1/10$, $p < .01$). These factors accounted for 6.8% and 2.4% of the variance in r_{est} error,

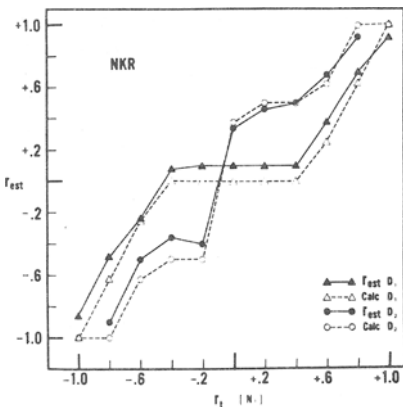


Fig. 2. Partitioned r_{est} functions for sequences of 20 events (N1) under the NKR condition. Functions are partitioned according to two levels of the parameter D.

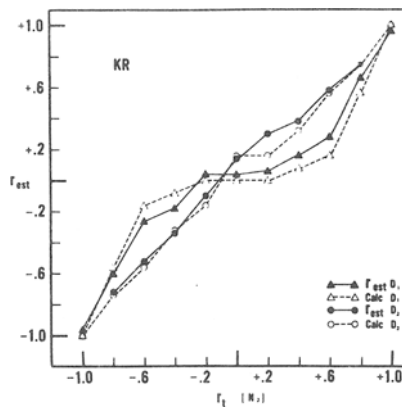


Fig. 3. Partitioned r_{est} functions for sequences of 30 events (N2) under the KR condition. Functions are partitioned according to two levels of the parameter D.

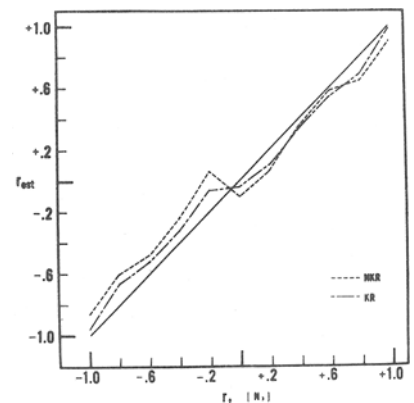


Fig. 5. Nonpartitioned r_{est} as a function of r_t for sequences of 40 events under NKR-KR conditions.

respectively. The sequence sign main effect is the result of higher r_{est} accuracy to positive sequences than to negative sequences and can be observed in Fig. 5. This result is consistent with that obtained by other investigators (e.g., Erlick & Mills, 1967; Jenkins & Ward, 1965). The main effect of N is due to a general increase in r_{est} accuracy as N increases. However, the largest increase in r_{est} accuracy was under N2.

A significant ND interaction ($F = 38.29$, $df = 2/20$, $p < .01$) can be attributed to the simple effects of N2D2 that produce a relative decrease in mean error (see Fig. 3, r_{est} D2). There were no effects across N for D1 sequences.

A significant KR-NKR main effect was not present and accounted for only 1.6% of total variance in r_{est} error. However, the three-way interaction of KR-NKR with D and Sequence Sign was borderline significant ($F = 7.26$, $df = 1/10$, $p < .05$) as was also the interaction between D and Sequence Sign ($F = 6.19$, $df = 1/10$, $p < .05$). Both of these effects are the result of the simple effects of negative sequences under the NKR condition. The effects are in terms of increased underestimation of D1, negative sequences with a larger underestimation error occurring under NKR rather than under KR.

Although it is apparent that the influence of D on r_{est} error was quite significant, the contribution of r_t relative to D and r_{est} was not specifically assessed in the analysis of variance. Linear regression coefficients were obtained for r_t , r_{est} under all conditions including nonpartitioned by D. These coefficients were all very high, ranging from .80 (NKR-N1D2) to .95 (KR-N3D2), indicating the effect of r_t on r_{est} was also significant. There were negligible differences among coefficients under KR-NKR and N conditions. However, as expected from the analysis of variance, differences did occur with respect to D. In order to evaluate the relative effectiveness of r_t and D as predictors of r_{est} , predictive efficiency ratios for r_{est} variance accounted for by D using the η coefficient to that accounted for by r_t using linear regression were determined. These ratios indicated that, with the exception of the KR-N2D2 condition, D is a better predictor of r_{est} than is r_t (range of ratios = 1.0 to 2.3, mean ratio = 1.4, $n = 12$).

In interpreting the results of this study, primary emphasis must be given to the observed significance of the relative-frequency parameter D, especially with regard to the effects of N.

It is suggested that r_{est} functions, when partitioned by D, may be highly specific

and predictable. Furthermore, functions of underestimation such as reported in a previous study of diagnostic performance based on perceptual quantification of relationship (Erlick & Mills, 1967) may be accounted for by a parameter similar to D. The functions in Fig. 4 support these statements in that r_{est} D1 functions reveal a tendency to underestimate r_t , while r_{est} D2 functions indicate a reciprocal tendency. Thus, specific independent D functions may result in r_{est} performance that underestimates r_t (e.g., see Fig. 2, calc D1N1) is nearly equivalent to r_t (e.g., see Fig. 3, calc D2N2) or overestimates r_t (e.g., see Fig. 2, calc D2N1).

This observation is important in understanding the N main effect and the ND interaction. Since there were no effects of D1 across N1, it may be concluded that D2 primarily under N2 is responsible for the observed reduction in r_{est} error. In examining the calculated D functions as N increases, it may be noted that D1 functions change little in shape and placement across N, while those of D2 change considerably from N1 to N2, from N2 to N3, and slightly from N1 to N3. Thus, in view of the strong relationship between D and r_{est} , it is not surprising that the variability of D2 functions across N result in corresponding changes in r_{est} . In addition, although calculated D1 values increase in magnitude with N, they generally remain proportionately constant, while those of D2 do not. The effect of N, then, is in terms of proportional changes in D across N rather than in terms of changes in magnitude. The fact that r_{est} performance is insensitive to changes in D magnitudes as long as there are no changes in proportion with increases in N is consistent with what would be expected on the basis of a relative-frequency principle.

Such a principle requires that relative frequencies of events are proportionately constant (i.e., their probabilities) across increases in sample size when sampling from the same population.

Therefore, it may be concluded that, under the conditions of this study, a relative-frequency parameter such as D can provide an appropriate model of Ss' diagnostic performance. Since the solution of D is based on obtaining the differential between frequencies of confirming and nonconfirming events, Ss' r_{est} performance appears to have been based on the perceived magnitude of this differential.

REFERENCES

- ERLICK, D. E., & MILLS, R. G. Perceptual quantification of conditional dependency. *Journal of Experimental Psychology*, 1967, 73, 9-14.
- INHOLDER, B., & PIAGET, J. *The growth of logical thinking from childhood to adolescence*. New York: Basic Books, 1958.
- JENKINS, H. M., & WARD, W. C. Judgment of contingency between responses and outcomes. *Psychological Monographs*, 1965, 79(1, Whole No. 594).
- SMEDSLUND, J. The concept of correlation in adults. *Scandinavian Journal of Psychology*, 1963, 4, 165-173.

NOTES

1. The research reported in this paper was sponsored by the Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson AFB, Ohio. This paper is identified as AMRL-TR-69-98. A more detailed analysis and discussion of this research relevant to human engineering interests may be found in AMRL-TR-68-135, "Use of contingent status information in diagnostic performance and related aspects for information design." Further reproduction is authorized to satisfy the needs of the United States government.

2. The author acknowledges the assistance of Beverly Hopkins, Chris Madigan, and John Wise in conducting this research effort. Also acknowledged is the critical help provided by Drs. D. E. Erlick and D. A. Topmiller.

Reaction to time-out and favorability of response alternatives

ROBERT H. WILLOUGHBY, *University of Massachusetts, Amherst, Mass. 01002*

Children's preference for one of two tasks was studied under three different conditions. In one condition, S received equal reinforcement (3:3) for responding on either task. In the remaining two

conditions, responses on one task produced reinforcement more frequently than did responses on the other at ratios of 3:2 and 3:1, respectively. Half of the Ss assigned to these three conditions received TOs on the more frequently reinforced task, with the remaining Ss serving as no-TO controls. Results revealed a significant tendency to