

Prediction of choice reaction time from information of individual stimuli¹

H. KAUFMAN, J. C. LAMB,² AND J. R. WALTER³

UNIVERSITY OF CONNECTICUT AND GENERAL DYNAMICS ELECTRIC BOAT DIVISION

A model is developed to predict the reaction time to stimuli using information content of the individual stimuli. Review of previous research in the area showed inconsistencies due possibly to the mode of stimulus presentation or to the number of stimuli. Two experiments designed to cover the range of variables used in previous experiments were conducted. The results indicate that the model makes accurate predictions over the entire range of variables used. A comparison to earlier studies also indicates a close correspondence.

Since shortly after Shannon (1948) proposed a mathematical theory of information and information transmission, psychologists have sought to apply this metric to human behavior. One part of this effort has been directed toward finding a general relationship between information content and speed of processing stimuli.

Hick (1952) found that choice reaction time (CRT) was an increasing linear function of the information transmitted with equally probable alternative stimuli (ELA), regardless of whether Ss made errors or not. This result has become known as Hick's law.

The function, $RT = a + b \log(n + 1)$, where n is the number of stimulus alternatives, has been called Hick's law (Welford, 1960). A more general formulation based on Hick's original results and later extensions (Hyman, 1953; Bricker, 1955; Lamb & Kaufman, 1965) is $RT = a + bT$, where T is transmitted information or, as stated by Hick, "... the rate of gain of information is, on the average, constant with respect to time ... [p. 11]."

In an attempt to extend Hick's law to the case of unequally probable alternative stimuli (ULA), Hyman (1953) reported an experiment using a situation similar to Hick's. His findings supported Hick's to the extent that CRT for the information in ELA stimuli was linear, CRT for the average information (H) in ULA stimuli was also linear, and the slopes of the two functions were not significantly different. However, CRT for the information in individual stimuli (h = surprisal value = individual stimulus information) deviated significantly from CRT values predicted by using the combined ELA and ULA average information curves. Hyman (1953)

suggested that "... different laws and equations [p. 194]" would be necessary to predict CRT to individual stimuli.

Lamb and Kaufman (1965) and Kaufman and Levy (1966) found results for ULA stimuli that differed in form from Hyman's, although they were still not consistent with the average information functions. All these experiments had used a discrimination rather than an absolute judgment situation for ULA stimuli. In these cases, Ss made essentially no errors, so that the function relating CRT to information was based on stimulus information rather than on transmitted information.⁴ Kaufman and Lamb (1966) reported an experiment using absolute judgments. Their results were similar to Hyman's, and, from the evidence, they postulated that the differences between ULA points and the ELA linear function are a function of S's threshold for differential stimulus probabilities. While their explanation could be applied, in a general way, to the previous studies, it did not make quantitative predictions. In addition, their experiment differed from previous studies in two ways. First, they used only two stimuli for all conditions in which stimuli were not equally probable and, second, they used an absolute judgment situation where other studies have used discriminative judgments.

The present study presents an extension of the reasoning advocated by Kaufman and Lamb (1966) to allow quantitative predictions and the results of two experiments designed to test the effects of the number of ULA stimuli and the difference between discrimination and absolute judgment on the predictions.

METHOD

Experimental Conditions

The same set of stimulus probabilities were used for both experiments. These conditions are presented in Table 1.

Apparatus

The same apparatus was used for both experiments. It consisted of a Gerbrands tachistoscope, a Gerbrands electronic voice key, and a Hunter millisecond timer. The S held a button in his hand that initiated a trial by presenting a white stimulus card in the tachistoscope and starting the timer. His vocal response terminated the trial via the voice key, stopping the timer and removing the stimulus from view. The use of S-initiated trials eliminated all uncertainty about the time of stimulus onset.

Hick made the explicit assumption that the uncertainty associated with the time of stimulus onset added exactly one additional alternative to the situation. In the present case, S initiates stimulus onset and eliminates the one alternative associated with stimulus onset uncertainty.

Subjects

The Ss were 112 male and female undergraduates at the University of Connecticut. Forty-eight were assigned randomly six to a group for Experiment 1, and the remainder were assigned eight to a group for Experiment 2.

Experiment 1

This experiment was designed to replicate Hyman's original experiment as closely as possible and served as the discriminative situation.

Table 1
Stimuli Number, Probability, and Information

Condition	Number Alternatives	Probability	Stimulus Information	Average Information
1) 2ELA	2	1/2 (.500)	1.0	1.0
2) 4ELA	4	1/4 (.250)	2.0	2.0
3) 8ELA	8	1/8 (.125)	3.0	3.0
4) 2ULA	1	7/8 (.875)	0.1926	0.5436
	1	1/8 (.125)	3.0	
5) 4ULA-High	1	7/8 (.875)	0.1926	0.7417
	3	1/24 (.042)	4.585	
6) 4ULA-Low	1	1/2 (.500)	1.0	1.7925
	3	1/6 (.167)	2.585	
7) 8ULA-High	1	7/8 (.875)	0.1926	0.8945
	7	1/56 (.018)	5.8074	
8) 8ULA-Low	1	1/2 (.500)	1.0	2.4037
	7	1/14 (.071)	3.8074	

Stimuli

Stimuli consisted of white cards with black stimuli, Xs and Os, 7/8 in. in size. The stimulus locations used were the four outermost corners and next inner four corners of an imaginary 6 by 6 matrix. *Bun, boo, bee, bore, bive, bix, bev, and bate* were the eight location names, of which two, four, or all were used, depending on the condition. Each side of the matrix made a visual angle of approximately 5 deg at S's location. The matrix was centered on the white card.

Procedure

Each S attended two sessions approximately 1 h in length. During the first (practice) session, E explained the nature of the experiment and described the operation of the equipment. On each trial, S viewed a display consisting entirely of Xs in the stimulus positions. After E signaled that the apparatus was ready, S pressed a hand-held button starting the timer and presenting another card identical to the one described above, except that one of the Xs was replaced by an O. The S's response, name of the location of the O, operated a voice key stopping the timer and returning the initial display into view.

The S was told the probabilities of an O appearing in each of the locations used in his condition; he was urged to be as fast and as accurate as possible. Then E presented 100 practice trials, with stimuli in a random order. During the second session, S was given 130 trials, of which approximately the last 100 were analyzed.

Experiment 2

This experiment followed closely the procedure used by Kaufman and Lamb (1966), serving as the absolute judgment condition.

Stimuli

The stimuli were eight 1/8-in-wide black lines on white backgrounds; their horizontal lengths varied from 1.00 to 4.04 in. by successive ratios of 1.22 to 1, which kept the discriminability between adjacent lines relatively constant over the entire range (Doherty, 1965). The response names used for the stimuli, in order of increasing length, were *bun, boo, bee, bore, bive, bix, bev, and bate*.

Procedure

Each S participated in two sessions of approximately 45 min. In the first session, S was given instructions, was shown the particular stimuli to be used, and performed 84 practice trials. In the second session, S was again shown the stimuli and was reminded of their particular frequency distribution. Then, after 28 practice trials, approximately 100 test trials were given.

On each trial, S pushed a hand-held button, a stimulus appeared on the tachistoscope, and the timer started. When S responded, the stimulus disappeared and the timer stopped. The E then recorded S's response and time and repeated the correct response to S.

RESULTS

Median CRT and transmitted information were calculated for each S for the second or test session for both experiments. Medians were used to minimize effects of skewness and heterogeneity in the data.

CRT for Average Information

CRT for ELA situations. For the first experiment (discriminative), the regression line for CRT vs $T(S;R)$ [RT = $.204 + .199 T(S;R)$, $r = .92$] was calculated; the line had a significant linear component ($F = 86.38$, $df = 1/16$, $p < .001$) when tested against deviations from linearity and no significant higher-order components. For the second experiment, CRT vs $T(S;R)$ [RT = $.110 + .373 T(S;R)$, $r = .91$] also had a significant linear effect ($F = 93.65$, $df = 1/19$, $p < .0005$) and no significant higher-order effect. This is consistent with Hick's original finding and was entirely expected.

CRT for ULA situations. CRT vs $T(S;R)$ [RT = $.243 + .180 T(S;R)$, $r = .89$, for Experiment 1 and RT = $.306 + .226 T(S;R)$, $r = .80$] was calculated for both experiments and was found to have a significant linear component ($p < .001$ for both experiments) and no higher-order effects. Again, this is the result found by Hyman (1953).

Comparison of CRT for ELA and ULA combined. The results stated so far indicate that the rate of gain of information in absolute judgments is a constant for ELA and a constant for ULA. For Hick's law to be more generally true, these constants should be identical. The ELA and ULA regressions of CRT on T should be the same.

Testing the regression coefficient and y-intercept for the first experiment, it was found that these were not significantly different for the ELA and ULA conditions. This indicates that the same function can account for both conditions, and the same process is assumed to be acting in both cases.

For the second experiment, however, the difference between the functions was significant ($F = 6.3$, $df = 2/59$, $p < .005$) for CRT vs $T(S;R)$. The difference between the functions was not significant for $T(S', R', S;R)$ (RT = $.323 + .190 T(S', R', S;R)$, $r = .78$).

A possible explanation of the significant

difference between the ULA and ELA functions for $T(S;R)$ is the existence of context effects in the ELA data. The amount of information being processed includes some information about the stimulus and the response on the previous trial, and, of course, it takes time to process it. The additional information processed is not included in the $T(S;R)$ measure but is in the $T(S', R', S;R)$ measure. The effect, especially in the 8 ELA condition, causes the observed CRT to be an overestimate of the time required to process the amount of information indicated by $T(S;R)$. This error leads to the apparently higher slope and, in general, the difference between the functions. Because of this sequential-dependency effect in the data, the function relating CRT to $T(S', R', S;R)$ should be taken as more representative of the information-processing behavior of the human making absolute judgments. The evidence suggests that the difference in transmission rate between discrimination and absolute judgment reflects the information S processes in making his judgments. In the discriminative situation, all the information processed is contained in the present stimulus; in the absolute judgment situation, some of the information is contained in the results of previous trials, i.e., context effects.

CRT for individual stimuli. There is no measure corresponding to $T(S;R)$ for individual stimuli that can take into account the information lost in processing (errors). As previous Es discovered, the relationship of CRT to h (surprisal) was not consistent with the relationship of CRT to $R(S;R)$ or to $H(S;R)$. These results were confirmed in both experiments of the present study. The general results were of the form found by Hyman (1953) and by Kaufman and Lamb (1966). The basic premise of Kaufman and Lamb (1966) was that S processed information sequentially depending on a threshold for differential probabilities. A quantitative extension of this hypothesis is explained below.

Sequential-Processing Model

The sequential-processing hypothesis maintains that, given a set of alternatives from which one is to be chosen, the human information processor makes successive binary choices between one alternative and those which remain. This sequence continues until a specific alternative is chosen. The information gained on each stage is a function of the relative frequencies of the different alternatives in the subset left after the previous stage, and, assuming no errors, the amount of information gained on any trial, averaged over many trials, is H , the average amount of information presented. Thus, minimum

CRT will occur on the trials that S starts the process with the alternative that actually occurred, that is, when the first stage is the choice between the one that occurred and the others. To have a low mean CRT to a given alternative, S must consistently start the process with it when it occurs. In the ULA case, one of the alternatives is "marked" by occurring more often than the others, and, to the extent that S can take advantage of the fact and prepare himself to start the process with the marked alternative, he can achieve significant reductions in CRT to it. In conditions of equally likely alternatives, the probability distribution marks no particular alternative, hence, the model assumes, S chooses the starting alternative randomly and therefore gives, roughly, equal average CRTs to all the alternatives in the set.

In general then, on every trial, S makes an initial decision as to whether the most probable stimulus has occurred. If it has, the S's reduction in uncertainty is equivalent to the information in the most probable stimulus plus the residual information in all remaining stimuli. Thus, for a set of stimuli, 1 to n, ranked in order of probability, the reduction in information processing for the most probable stimulus is

$$\hat{h}_1 = [-p_1 \log_2 p_1] - [(1 - p_1) \log_2 (1 - p_1)]. \quad (1)$$

If the most probable stimulus does not occur, then the time required for this decision is the time that S uses to process the information in Eq. 1. Next, S decides if the second most probable stimulus has occurred. The total reduction for the second most probable stimulus occurring is

$$\hat{h}_2 = \hat{h}_1 - \frac{p_2}{(1 - p_1)} \log_2 \left[\frac{p_2}{(1 - p_1)} \right] - \left[1 - \frac{p_2}{(1 - p_1)} \right] \log_2 \left[1 - \frac{p_2}{(1 - p_1)} \right]$$

or first decision plus second stimulus decision plus residual information. This process is repeated until a decision has been made for all stimuli.

It can be shown that, if at any point the remaining stimuli are all equally probable, the residual term is simply \log_2 of the number of stimuli remaining. Thus, for the present experiment, two equations are sufficient, Eq. 1 for the most probable stimulus, and

$$\hat{h}_{2-N} = -\hat{h}_1 + \log_2 (n - 1) \quad (2)$$

for all other stimuli.

Table 2
Predicted and Actual Results

Condition	Probability	n	h (Effective Surprisal)	Absolute Judgment		Discriminative Situation Predicted	Actual
				Predicted	Actual		
2-	.875	1	0.5436	.410	.342*	.332	.315
	.125	1	0.5436	.410	.416	.332	.370
4-Low	.500	1	1.000	.507	.556	.335	.440
	.167	3	2.585	.843	.655*	.654	.630
4-High	.875	1	0.5436	.410	.431	.332	.314
	.042	3	2.129	.746	.642	.562	.572
8-Low	.500	1	1.0	.507	.528	.335	.397
	.071	7	3.807	1.102	.876*	.900	.786
8-High	.875	1	0.5436	.410	.445	.332	.324
	.018	7	3.351	1.005	.891	.808	.937

* Significant $p < .05$

Results of Predictions

The reduction in uncertainty (h = effective surprisal) was calculated for each stimulus on the basis of Eqs. 1 and 2, and the CRT to that information was estimated. Estimates were based on CRT vs $T(S;R)$ for the discriminative situation and CRT vs $T(S', R', S;R)$ for the absolute-judgment situation. The function used for this prediction was that of the combined ELA and ULA average information vs CRT. The results for predicted and actual CRT for both conditions are shown in Table 2. Differences between the predicted and mean CRT for each stimulus were evaluated using t tests. Of the 20 analyses performed, only 3 were significant ($p < .05$).

The average error was greater for absolute judgment than for discrimination condition, and all significant ts occurred for this situation. This probably reflects the context effects noted earlier and also the difficulty, in an absolute-judgment task, that S has in establishing a clear perceptual standard for the marked alternative at the start of each trial.

In general, the model handles equally well each number-of-alternative condition. One noticeable, if small, discrepancy appears in the two-alternative case, where RT to the less frequent stimulus is consistently overestimated by the model.

Some sort of threshold process, as earlier suggested by the present authors, seems to operate against the predicted effects in the two-alternatives case but not for larger number of alternatives.

In order to test further the generality of the model, predictions were made from the available data points given by Hyman (1953) and by Stone and Callaway (1964). The results are shown in Table 3. While no statistical tests were possible, since individual S's scores were not given, the agreement between predicted and average values is about the same as the present results. The second part of the Stone and Callaway data is for a three-stage processing situation, a situation not tested in the present study.

DISCUSSION

The effectiveness of the proposed model can be seen from the fact that it does make accurate quantitative predictions using only the information available in individual stimuli. The main accomplishment of the model is that it extends the domain of Hick's law (constant rate of information processed), at least within the choice-reaction-time situation, without invoking any special assumptions about decision thresholds or subjective probabilities. Furthermore, with the two parameters varied in this experiment (discrimination-identification and number

Table 3
Predicted and Actual RTs from Previous Experiments

Probability	N	Total N	Reduction in Uncertainty	Predicted	Actual
Hyman (1953)					
.813	1	4	0.69	.317	.306
.062	3	4	2.28	.560	.585
Stone & Callaway (1964)					
.812	1	4	0.69	.326	.325
.062	3	4	2.28	.388	.375
.500	1	4	1.0	.338	.345
.250	1	4	2.0	.377	.370
.125	2	4	3.0	.416	.375

of ULA alternatives), covering the range of previous studies, the model has been found applicable to a variety of ULA conditions. Any generality beyond this must be determined empirically.

It might be argued that some function relating CRT to information for both ELA and ULA could be found that would give more accurate predictions than the pooled ULA and ELA information used in this study. Our purpose in the present research, however, was to show that at least one such function exists: namely, the line fit to the combined ELA and ULA (average) data.

In summary, the present study shows that accurate predictions of CRT for a variety of both ELA and ULA situations can be made using informational concepts that are consistent with the principle that information is processed at a constant rate. Additional tests of the extension to three- or four-stage situations are planned.

REFERENCES

BRICKER, P. D. Information measurement and

- reaction time. In H. Quastler (Ed.), *Information theory in psychology*. Glencoe, Ill.: Free Press, 1955. Pp. 350-359.
- DOHERTY, M. Information and discriminability as determinants of absolute judgment choice reaction time. Unpublished doctoral dissertation, University of Connecticut, 1965.
- HICK, W. E. On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 1952, 4, 11-26.
- HYMAN, R. Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, 1953, 45, 188-196.
- KAUFMAN, H., & LAMB, J. Choice reaction time and unequal stimulus frequencies in an absolute judgment situation. *Perception & Psychophysics*, 1966, 1, 354-387.
- KAUFMAN, H., & LEVY, R. A further test of Hick's law with unequally likely alternatives. *Perceptual & Motor Skills*, 1966, 22, 967-970.
- LAMB, J., & KAUFMAN, H. Information transmission with unequally likely alternatives. *Perceptual & Motor Skills*, 1965, 21, 255-259.
- MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, 63, 81-97.
- SHANNON, C. E. Mathematical theory of communication. *Bell System Technical Journal*, 1948, 27, 379.
- STONE, G. C., & CALLOWAY, E. Effects of stimulus probability on reaction time in a

number naming task. *Quarterly Journal of Experimental Psychology*, 1964, 16, 47-55.

WELFORD, A. T. The measurement of sensory motor performance: Survey and reappraisal of twelve years' progress. *Ergonomics*, 1960, 3, 189-229.

NOTES

1. Acknowledgment is due the Office of Naval Research which supported this research through a prime contract, NONr 2512(00), with General Dynamics Electric Boat Division as a part of the SUBIC (SUBmarine Integrated Control) program.
 2. Address: General Dynamics Electric Boat Division, Eastern Point Road, Groton, Connecticut 06340.
 3. Now at Bethlehem Steel Company.
 4. The two are identical where no errors are made.
 5. $T(S;R)$ is the information transmitted from the stimulus on the present trial to the present response. $T(S', R', S;R)$ is the information transmitted from the stimulus and the response on the previous trial and the present stimulus.
- (Accepted for publication June 19, 1969.)