On the trace strength of responses varying in correctness

ALEXANDER J. WEARING*
Yale University, New Haven, Conn. 06510

A paired-associate learning experiment was carried out to see if two presumed measures of memory trace strength, subjective certainty and latency, were able to distinguish between responses varying in their degree of correctness. Evidence was found for three discrete incorrect-response states.

The use of subjective verbal reports of either confidence judgments or mnemonics in verbal learning studies is not widespread. But evidence is accumulating that Ss' reports of mnemonic utilization are correlated with performance on acquisition and retention tests (Adams, 1967; Montague, Adams, & Kiess, 1966; Montague & Wearing, 1967; Wearing, Walker, & Montague, 1967) and that Ss can make reliable judgments about the correctness, or otherwise, of their responses (Underwood & Freund, 1968).

Indeed, studies now exist which suggest that Ss can judge the strength of the memory trace underlying a response, even though they cannot produce the response itself. Examples are the "tip-of-the-tongue" (TOT) phenomenon (Brown & McNeill, 1966) and the "feeling-of-knowing" (FK) experience described by Hart (1965, 1967).

These findings suggest that the strength of the memory trace varies continuously (the TOT or FK phenomena could not occur otherwise) and that, to some extent, S is aware of changes in trace strength, i.e., he can make reliable judgments about it. If confidence ratings do vary with trace strength, then one would also expect a strong relationship between subjective certainty and latency, since latency is a generally accepted measure of memory trace strength (Adams, 1967). In particular, this relationship should hold when the range of trace strength variation is restricted to incorrect responses, if both the Brown and McNeill and the Hart findings are to be confirmed.

An experiment was carried out in which latency measures and confidence judgments were obtained on both correct and incorrect responses. The responses themselves were scored on a degree of

correctness scale, which enabled the relationships between subjective certainty, response latency, and response correctness to be investigated.

SUBJECTS

Sixty-one paid female undergraduate students served as Ss.

VERBAL MATERIALS

Sixty pairs of CVCs were used, ranging evenly from low to high AV (Archer, 1960).

PROCEDURE

Ss were run in groups on the University of Illinois PLATO system (Bitzer, Hicks, Johnson, & Lyman, 1967). Instructions, stimuli, and responses appeared on a CRT in front of each S, and Ss typed their responses on the keyset. To familiarize Ss with computer-controlled learning, a practice list of eight letter pairs was learned.

Ss were then instructed that 60 CVC pairs would be presented at a rate of 15 sec per trial for one trial, during which they were to learn as many pairs as possible. A bonus of 2 cents per correct recall was offered as an incentive to learn. Immediately after acquisition, a test trial was given, with each stimulus showing for 15 sec.

Twenty-four hours later, Ss were given a second retention test, during which they were given 15 sec to recall the response member of each pair and to indicate on a 5-point scale the degree of certainty they felt in the correctness of their response. To reduce variability in the data due to individual differences in typing skill, Ss were required to type only with their preferred hands, hunt-and-peck fashion.

RESULTS

The responses were scored in the following way. If a letter from the CVC was correct, it was given one point. If, in addition, the letter was in the correct position, it was given a second point. Thus, the maximum possible score was six points (obtainable only if a response was correct) and the minimum was zero. In order to distinguish errors of commission, where none of the letters were correct, from

errors of omission, an extra point was given for any actual response.

The relationship between confidence ratings, latencies, and degrees of correctness is shown in Table 1. It seems true that faster response latencies and high levels of subjective confidence were associated with more nearly correct responses. There were no significant differences between the subjective certainty means for Categories 2, 3, 4, and 5, although Categories 1, 6, and 7 were significantly different from each other and all other categories. This pattern of an initial rise in certainty, followed by a plateau, and succeeded in turn by a sharp increase in confidence was paralleled in the latency figures. It looks as though only about three noncorrect response states can be distinguished: no response (omission), near-random guess (Categories 2, 3, 4, and 5), and TOT (Category 6).

DISCUSSION

These results do not support an all-or-none theory of learning, as Ss apparently can make distinctions between wrong responses which vary in their degree of correctness. That Ss' judgments are related to trace strength is shown by the finding that response latency (a measure of trace strength) was closely related to subjective certainty.

However, a simple incremental model will not handle these data either. There is a sharp change in both latencies and confidence ratings in moving from response Category 6 (nearly correct) to response Category 7. It appears as though there might be a discontinuity in trace strength at the point at which a response becomes correct, as well as prior to that stage. A further reason why the simple incremental model will not do is that the relationship between degrees of correctness and latency was not linear, a result which confirmed that of an earlier experiment (reported in part by Wearing, Walker, & Montague, 1967), in which the latencies for Response Categories 2-7 were 6.09, 4.81, 5.18, 5.18, 4.10, and 2.62 sec, respectively.

Table 1
Mean Confidence Ratings and Mean Latencies
as a Function of Response Correctness

Degree of Correctness	Number of Cases Per Category	Mean Confidence Ratings	Mean Latencies (in Seconds)
1 (Omission)	741	1.4	
2	1175	1.8	5.18
3	134	2.1	5.02
4	458	2.0	5.02
5	37	2.2	4.74
6	126	2.7	3.81
7 (Correct)	1088	4.2	2.58

*The author gratefully acknowledges the assistance of Paula Groene, Linda Kalmanek, and Rosemary Wearing in the collection of these data, and the comments and suggestions of William Montague and Clinton Walker on an earlier draft of this paper.

What these data do indicate is that the growth function of trace strength may not be as simple as is currently believed or, alternatively, that the relationship between different indices of trace strength (degree of correctness, subjective certainty, and response latency) is more complex than has as yet been imagined.

REFERENCES

- ADAMS, J. A. *Human memory*. New York: McGraw-Hill, 1967.
- ARCHER, E. J. A re-evaluation of the meaningfulness of all possible CVC trigrams. *Psychological Monographs*. 1960, 74(10, Whole No. 497).
- BITZER, D. L., HICKS, B., JOHNSON, R., & HYMAN, E. R. The PLATO system: Current research and development. *IEEE Transactions on Human Factors in Electronics*, 1967, 8, 64-70.
- BROWN, R., & McNEILL, D. The "tip of the tongue" phenomenon. *Journal of Verbal Learning & Verbal Behavior*, 1966, 5, 325-337.
- HART, J. T. Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, 1965, 56, 208-218.
- HART, J. T. Second-try recall, recognition, and the memory-monitoring process. *Journal of Educational Psychology*, 1967, 58, 193-197.
- MONTAGUE, W. E., ADAMS, J. A., & KIESS, H. O. Forgetting and natural language mediation. *Journal of Experimental Psychology*, 1966, 72, 829-833.
- MONTAGUE, W. E., & WEARING, A. J. The complexity of natural language mediators and its relation to paired-associate learning. *Psychonomic Science*, 1967, 7, 135-136.
- UNDERWOOD, B., & FREUND, J. Errors in recognition learning and retention. *Journal of Experimental Psychology*, 1968, 78, 55-63.
- WEARING, A. J., WALKER, C. B., & MONTAGUE, W. E. Recall of paired associates as a function of their associability. *Psychonomic Science*, 1967, 9, 533-534.

Probability and expectancy in two-choice reaction time*

JAMES V. HINRICHS

University of Iowa, Iowa City, Iowa 52240

RTs were measured to two stimuli which were presented in a 2 : 1 ratio. The mean RT was inversely related to the probability of stimulus presentation, but no difference was found when the mean RT was calculated conditional on Ss' predictions of stimulus presentation. The probability effect in choice RT was interpreted as depending upon a weighted combination of fast RTs to correctly predicted stimuli and slow RTs to incorrectly predicted stimuli.

Two recent theories of choice reaction time (RT) have been proposed which use expectancy concepts (i.e., states of preparedness or subjective probability) to account for various effects commonly found in choice RT experiments (Falmagne, 1965; Laming, 1969). In these theories the expectancy state is hypothetical and is not directly observable. One possible way to assess the expectancy of the S is simply to ask him to predict which stimulus will be presented next.

Several experiments have demonstrated

*The assistance of S. M. Holtkamp and P. L. Krainz is gratefully acknowledged.

that the RT to correctly anticipated stimuli is shorter than to incorrectly anticipated stimuli (e.g., Bernstein & Reese, 1965). Hinrichs & Krainz (1970) have shown that the difference in RT to correctly and incorrectly predicted stimuli is due to a stimulus-expectancy effect rather than to differential preparation to execute the associated responses. The purpose of the present study was to extend the application of expectancy effects to stimulus probability effects in two-choice RT experiments.

The probability effect, also called the frequency or uncertainty effect, refers to

the relationship between RT and the probability of stimulus occurrence: as the probability of a stimulus increases, the RT to that stimulus decreases (see Smith, 1968, for a recent review). Because Ss tend to match their predictions of stimulus presentations to the a priori probability of occurrence (e.g., Estes, 1964), a possible interpretation of the probability effect in two-choice RT is that the difference in RT is a function of the probability of a correct anticipation. As the probability of occurrence increases, the probability of predicting that stimulus also increases and, correspondingly, the probability of a correct prediction increases. Therefore, the difference in mean RT to stimuli with different frequency might be attributed to the relative proportion of RTs to correctly and incorrectly predicted stimulus occurrences.

Bernstein & Reese (1965) have already shown, in an experiment which varied stimulus uncertainty across blocks of trials, that when RTs are partitioned into correct and incorrect predictions, a frequency effect is found only for incorrectly predicted stimuli. Across levels of stimulus uncertainty, as the stimulus information increased, the mean RT to incorrectly predicted stimuli increased; the mean RT to correctly predicted stimuli did not change. In the present experiment, the probability effect was examined within a single probability level rather than across levels. The primary goal of the study was to assess the adequacy of stimulus predictions as a measure of subjective expectancies and to examine the extent to which the frequency effect in two-choice RT may be explained as the weighted combination of expected and unexpected stimulus presentations.

SUBJECTS

Twenty-four University of Iowa undergraduates, 12 males and 12 females, served as Ss, fulfilling an introductory psychology course requirement. The Ss were randomly assigned to two groups, each group containing six males and six females.

MATERIALS AND APPARATUS

The stimuli were the digits 1 and 2, presented by an Industrial Electronic Engineers Bina-View self-decoding display cell (Model KA-12/12-093-E-1886) which has a 5.8 x 4.6 cm display screen. Each S was tested on eight blocks of 30 trials per block. Within each block one stimulus occurred twice as often as the other. The assignment of 1 or 2 as the more-frequent (MF) or less-frequent (LF) stimulus was counterbalanced across Ss but maintained across all eight blocks of trials for each S.

The presentation of the stimuli was controlled by prepunched paper tapes.