Memory & Cognition
1973, Vol. 1, No. 3, 383-386

Mental set and mental arithmetic*

IRVING BIEDERMAN†
State University of New York, Buffalo, New York 14226

Ss performed mental arithmetic problems in which they added, subtracted, or multiplied two one-digit numbers. The presentation order of the operator symbol and the digits was varied. With three possible operators, presentation of the operator prior to the digits (OD) led to faster RTs. With two possible operators, the opposite order (digits prior to operator, DO) led to faster RTs, because RTs in the OD condition were unaffected by the number of possible operators. These results are discussed in terms of the tradeoff between accessing active memory for a small number of items in the DO condition vs retrieving information from relatively large tables in long-term memory in the OD condition.

Rational analysis of complex mental activity often leads to the positing of internal structures that are said to be hierarchically organized in the sense that the processing of specific information (or signals, exemplars, or data) is determined by the activation of a rule (or concept, category, strategy, operator, or set) (Hebb, 1966; Lashley, 1951; Miller, Galanter & Pribram, 1969; Reitman, 1965).

Several partial advance information experiments have reported that, when the processing of a signal is contingent upon the specification of a rule, it is more efficient to present the rule prior to the signal than in the reverse order (Bernstein & Segal, 1969; Davis, 1967; Shaffer, 1965; Whitman & Geller, 1971). In many of these experiments, however, the designation of rule and signal components of the stimulus complex was arbitrary—the signals conveyed as much information (in the sense of eliminating alternative responses) as the rule. (See Biederman, 1972, p. 225, for a more extensive discussion of this problem.) By the very nature of hierarchically organized memory structures in which there is a one-to-many divergence as one proceeds from the general to the specific, rules will typically convey *less* response information than the signals. Consider the case of mental arithmetic, the task of interest in the present experiment. In adding, subtracting, or multiplying two one-digit numbers (from 1 to 9), presentation of the two numbers prior to the operator will leave only three possible responses. For example, if the digits 7 and 3 are presented, S can retrieve the possible answers and store them as the response terms in three paired associates (+ → 10, – → 4, and X → 21) in active memory (STM). However, in a condition where operators are presented prior to the digits, if a "+" is presented there are 16 possible answers, for "X" there are 35, and for "–" there are 9. On the average, an operator by itself will reduce the set of possible answers from 45 to 21—still far too many for immediate memory.[1] Assuming a sufficient interval with which to retrieve (or compute) the possible responses, if response uncertainty was a potent determiner of performance, then a presentation mode in which digits were presented prior to the operator would be favored.

However, if we consider the typical manner in which mental arithmetic is performed, it is clearly the case that presenting the operator prior to the digits is a condition that is more compatible with our experience of generally knowing what we will be doing with numbers prior to seeing them. Whether mental arithmetic is a table look-up (or retrieval) process or a rapid computational (or constructive) process (Groen & Parkman, 1972), considerations of typical mental arithmetic performance—and, thus, the organization of arithmetic information in long-term memory (LTM)—would appear to favor a presentation mode in which operators are presented first.

Thus, the present experiment, in which the order of presentation of the operator and digit was varied, was a case study of how well-learned retrieval or computational routines of relatively high response uncertainty fare when pitted against the utilization of active memory for a few items. The time for accessing STM is a direct function of the number of items to be scanned, while accessing LTM through well-learned categories is barely affected by category size, if it is affected at all. Thus, estimates of the scanning rate for unfamiliar sets of familiar items in immediate memory are in the order of 35-40 msec/item (Sternberg, 1966, 1967), but categorizing a word as a living thing requires only 8 msec longer than categorizing it as an animal (Landaur & Freedman, 1968). These results are compatible with the thinking that immediate memory is sequentially scanned (Sternberg, 1966) but information in LTM can be accessed in parallel (Landaur & Freedman, 1968; Rips, Shobin, & Smith, 1973; Theios, in press). For very high compatibility tasks, such as naming, accessing of memory can be described as content addressable: The response is stored at the same

*The author thanks Walter Cohen, A. L. Kamlet, E. E. Smith, and H. B. G. Thomas for their helpful suggestions. The research was supported by the Advanced Research Projects Agency, U.S. Department of Defense, and was monitored by the Air Force Office of Scientific Research under Contract AF 49(628)1235 with the Human Performance Center, University of Michigan.

†Requests for reprints should be sent to Irving Biederman, Department of Psychology, State University of New York, 4230 Ridge Lea Road, Buffalo, New York 14226.

location as the output of the pattern recognition process, so that no further scanning of memory is required to find the response (Greenwald, 1972; Theios, in press). Such high compatibility tasks reveal little or no effect of event uncertainty. If presenting the digits prior to the operator induces S to use a STM scan to do mental arithmetic, then that condition should show an effect of the number of possible answers. When the operator is presented prior to the digits, well-mastered routines for accessing LTM should be available with little or no effect of task uncertainty.

METHOD

Subjects

The Ss were three graduate students in philosophy and an undergraduate who were paid for their participation.

Apparatus and Procedure

S's task was to add, subtract, or multiply two one-digit numbers. Problems were displayed on a Nixie tube, horizontally centered between two in-line displays.

On each trial an arithmetic operator, "+," "X," or "–," was presented on the Nixie tube and a digit, 1-9, was presented on each of the in-line displays. The S would bark his answer into a voice key, which would stop a clock that had been initiated by the onset of the display. For the subtraction problems, S was instructed to give only the absolute difference. Voltages across the three display units were adjusted so as to yield equal rise times. The display subtended a visual angle of approximately 4 deg. E manually programmed the display and verbally provided speed and accuracy feedback after each response. S wore earphones through which the feedback information was provided and through which white noise was played at other times to mask distracting sounds.

Design

Each S participated in eight sessions. Each session consisted of 20 blocks of 30 trials (or problems). In half the blocks, those with operator set size of 2 (OP = 2), only two of the operators were presented, i.e., + and X, + and –, or X and –. In the other half, those with operator set size of 3 (OP = 3), all three operators were possible. Within each set size condition, on one of the blocks of trials both the operator and digits were presented simultaneously (SIM condition). The remaining nine blocks of trials were divided into three in which the operator preceded the digit (OD condition), three in which the digit preceded the operator (DO condition), and three in which the order of the operator and digits would vary from trial to trial (UN, for uncertain, condition) in random appearing fashion. The interstimulus interval between the presentation of the first and second kinds of information was .23, .57, or .94 sec, with one interval remaining constant for each of the three DO, OD, and UN blocks in each session. The order of administration of the conditions was balanced within each day, except that one of the SIM blocks was always administered first for warm-up. The SIM condition was, therefore, not directlycomparable to the other conditions. A given S had the same pair of operators in the OP = 2 condition throughout his participation in the experiment. Two of the Ss had + and –, one S had + and X, and the remaining S had X and –. Prior to performing each block, S was fully instructed as to the condition and allowed to perform several practice trials. A card describing the condition remained in S's view throughout the trial block.

RESULTS

There was no effect of introducing uncertainty about the ordering of operators and digits. Mean RT for the UN condition was .64 sec, which closely matched the .63-sec mean RT for the combined DO and OD conditions. Overall error rates were 8.4%, 8.9%, and 9.0% for the UN, DO, and OD conditions, respectively. When decomposed into DO and OD trials, the UN condition yielded results that were similar to the DO and OD conditions. These UN data will not, therefore, be discussed further.

Figure 1 shows the mean correct RTs as a function of order, interstimulus interval, and number of possible operators. The data for the individual Ss were highly similar to the composite shown in Fig. 1. There was a relatively large 76-msec effect of the number of possible operators in the DO condition but only a small 12-msec effect in the OD condition. This produced the most salient feature of these data, which is the interaction between order and number of possible operators. With OP = 3, OD led to shorter RTs than DO, but the reverse was true in the OP = 2 condition. In all conditions there was a consistent reduction in RTs from the 0- to the .57-sec ISI. As with most other partial advance information situations, the magnitude of the reduction in RTs was far less than the ISI. That is, the .23-sec ISI did not yield RTs that were .23 sec faster than the SIM condition. Learning reduced RTs from an overall mean of .74 sec on Session 1 to .57 sec on Session 8.

Separate analyses of variance were performed on the data from the two operator sets. In the OP = 3 condition, significant effects were found for order $[F(1,3) = 23.79, p < .05]$, ISI $[F(2,6) = 109.05, p < .01]$, and sessions $[F(7,21) = 20.53, p < .001]$. The only interaction that was significant was Order by Sessions $[F(7,21) = 3.05, p < .05]$; the advantage of the OD condition was reduced with practice. In the OP = 2 condition, significant effects were found for order $[F(1,3) = 44.15, p < .01]$, ISI $[F(2,6) = 11.60, p < .01]$, and sessions $[F(7,21) = 40.54, p < .001]$. The RTs in the DO = 2 condition steadily declined with longer ISIs, while the RTs in the OD = 2 condition remained relatively stable over the different ISIs. This produced a significant interaction between ISI and order $[F(2,6) = 27.99, p < .05]$. The F ratios for all the remaining interactions were approximately equal to 1.00.

For each block of 30 trials, a Pearson product-moment correlation coefficient was calculated between the RT to each problem and the measure of problem difficulty empirically derived by Thomas (1963). The measure, Q, is equal to log $(D_1 + D_2 + \text{Ans})$, where D_1 and D_2 are the digits of the problem and Ans is the answer to the problem. The paper and pencil tasks under which Q was derived were designed to reflect typical mental arithmetic performance. Q should be a predictor of RTs when the problems are solved by LTM retrieval or computational procedures. In the present experiment, the correlation coefficients between Q values and RTs were uniformly positive (mean Fisher z value of .292) in the OD condition but hovered about zero (.058) in the DO condition. Of the 192 DO-OD comparisons (three intervals by two operator set sizes by eight blocks by four Ss), 141 rs were more positive in the OD condition, 45 were more positive in the DO condition, and 6 were equal.

The RT variability was consistently greater in the DO than in the OD condition. The mean pooled (over Ss and sessions) standard deviation was .167 sec in the DO condition and .120 sec in the OD condition. Of the 192 standard deviation comparisons, 157 were larger in the DO condition, 29 were larger in the OD condition, and in 6 DO-OD comparisons the standard deviations were equal.

DISCUSSION

The different presentation orders induced different ways of performing the task. In the DO condition, S would determine the answers and store them as paired-associate response terms (to operator-symbol stimuli) in active memory. Consistent with this view is the 76-msec effect of the number of possible operators, a value remarkably close to what would be expected from the scanning of two additional items (an operator symbol and an answer) in immediate memory at a 35-40 msec/item scanning rate (Sternberg, 1966, 1967). Also consistent with anticipation in the DO condition is the lack of any correlation between RTs and Thomas's measure of problem difficulty. If the problem had already been looked up (or calculated), then there would be no relation of RTs with the difficulty of the look-up.

However, other considerations suggest that not all of the effect of the number of possible operators in the DO condition should be attributed to memory scanning. The anticipation of the operator during the ISI was often wrong and, from Ss' postexperimental protocols, often disruptive. From S F.K.: "The 1-sec DO often produced problems because it gave me too much time to think, i.e., time for me to read the two digits as either addition, subtraction, or multiplication [to already take one of these sets with respect to them before the operator came on] so that if the sign [operator] was not the one I had [semiawarely] taken to combine them with, I tended to be slow in responding." S A.L.: "The short intervals were easiest, as they provided enough time to jog the thinking machine to do a mental process without prejudging the operation. 'Digits' first was slow, as preguessing the operator was prevalent." S B.J.: "In both 1-sec DO conditions, I also had time to attempt concentration on the two or three alternatives, although [the] +/− condition allowed the most successes. +/X sometimes caught me off balance." Presumably, in the DO-2 condition, S would more often be able to complete his precalculations.

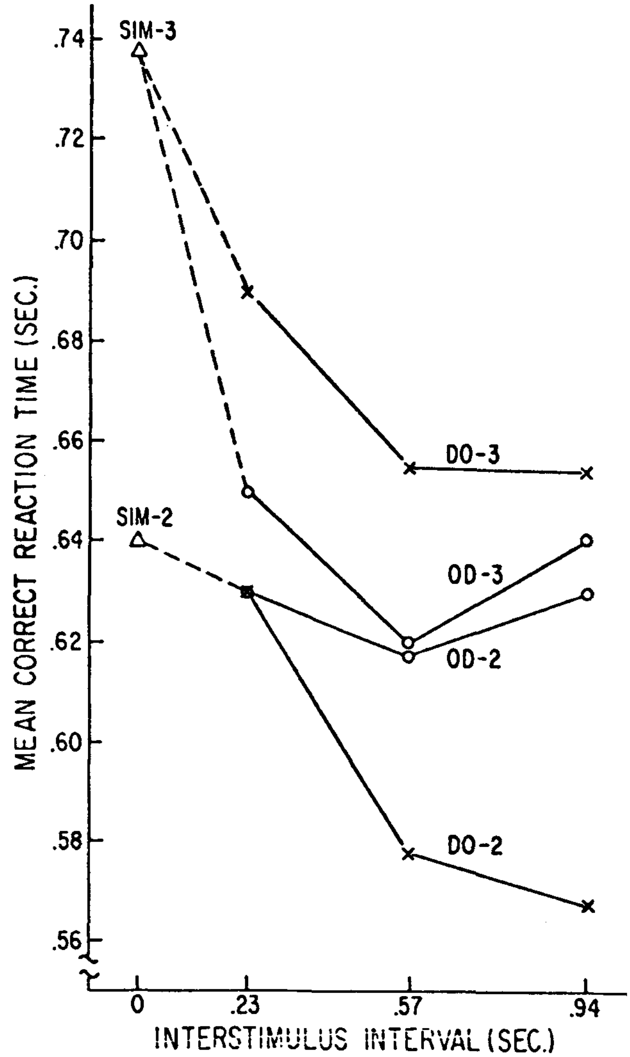

Fig. 1. Mean correct reaction time as a function of interstimulus interval, presentation order (DO = digits before operator, OD = operator before digits), and operator set size (2 or 3).

The experiment did not examine whether the disruption from an incorrect anticipation was due to the incompatibility of immediate memory scanning with a table look-up process (which was described in the protocols as "automatic" or "instinctive") or to response incompatibility between the anticipated number and the correct number. In any event, the larger RT variance associated with the DO condition compared to the OD condition—even in the DO-2 condition, where RTs were less than in the OD-2 condition—is consistent with this picture of a dual process in the DO condition.

The evidence that there may be a detrimental effect of partial advance information is not without precedent. Reicher (1969), in a letter identification task, found that giving S knowledge of the response alternatives (two letters) before viewing a briefly presented stimulus (a word, quadrigram, or letter) resulted in *worse* performance than when the response alternatives were made available only after S had viewed the stimulus. It is possible that the response alternatives enticed S to give up an efficient and automated mode of reading the display for some unautomated mode that would employ the response alternatives. A similar phenomenon might have been acting to the detriment of the DO-3 condition.

In the OD condition, there was much less tendency to anticipate a response. In fact, after the operator was identified, there was little to do but wait for the digits. Any effect of operator set size would largely be eliminated after a brief interval and this is what the data show. By .57 sec, the OD-3 condition is identical to the OD-2 condition. The rise in RTs between the .57- and .94-sec conditions could reflect some waning of attention during the relatively long interval.

REFERENCES

Bernstein, I., & Segal, E. M. Set and temporal integration. Perception & Psychophysics, 1968, 4, 233-236.

Biederman, I. Human performance in contingent information-processing tasks. Journal of Experimental Psychology, 1972, 93, 219-238.

Davis, R. Intermittency and selective attention. Acta Psychologica, 1967, 27, 57-63.

Greenwald, A. G. On doing two things at once: Time sharing as a function of ideomotor compatibility. Journal of Experimental Psychology, 1972, 94, 52-57.

Groen, G. J., & Parkman, J. M. A chronometric analysis of simple addition. Psychological Review, 1972, 79, 329-343.

Hebb, D. O. *A textbook of psychology.* 2nd ed. Philadelphia: Saunders, 1966.

Lashley, K. S. The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior.* New York: Wiley, 1951. Pp. 112-136.

Landaur, T. K., & Freedman, J. L. Information retrieval from long-term memory: Category size and recognition time. Journal of Verbal Learning & Verbal Behavior, 1968, 7, 291-295.

Miller, G. A., Galanter, E., & Pribram, K. H. *Plans and the structure of behavior.* New York: Holt, 1960.

Reicher, G. M. Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology, 1969, 81, 275-280.

Reitman, W. R. *Cognition and thought: An information-processing approach.* New York: Wiley, 1965.

Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. Journal of Verbal Learning & Verbal Behavior, 1973, 12, 1-20.

Shaffer, L. H. Choice reaction time with variable S-R mapping. Journal of Experimental Psychology, 1965, 70, 284-288.

Sternberg, S. High-speed scanning in human memory. Science, 1966, 153, 652-654.

Sternberg, S. Two operations in character recognition: Some evidence from reaction-time measurements. Perception & Psychophysics, 1967, 2, 45-53.

Theios, J. Reaction-time measurements in the study of memory processes: Theory and data. In G. H. Bower (Ed.), *The psychology of learning and motivation.* Vol. VII. New York: Academic Press, in press.

Thomas, H. B. G. Communication theory and the constellation hypothesis of calculation. Quarterly Journal of Experimental Psychology, 1963, 15, 173-191.

Whitman, C. P., & Geller, E. S. Stimulus anticipation in choice reaction time with variable S-R mapping. Journal of Experimental Psychology, 1972, 93, 433-434.

NOTE

1. Since the responses are not equally probable, there are 3.91, 2.96, and 5.09 bits of response uncertainty following presentation of the addition, subtraction, and multiplication operators, respectively.

(Received for publication February 18, 1973; accepted March 18, 1973.)