

SOME COUNTING MODELS FOR FIRST-GRADE PERFORMANCE DATA ON  
SIMPLE ADDITION FACTS

by

Patrick Suppes and Guy Groen

TECHNICAL REPORT NO. 90

January 14, 1966

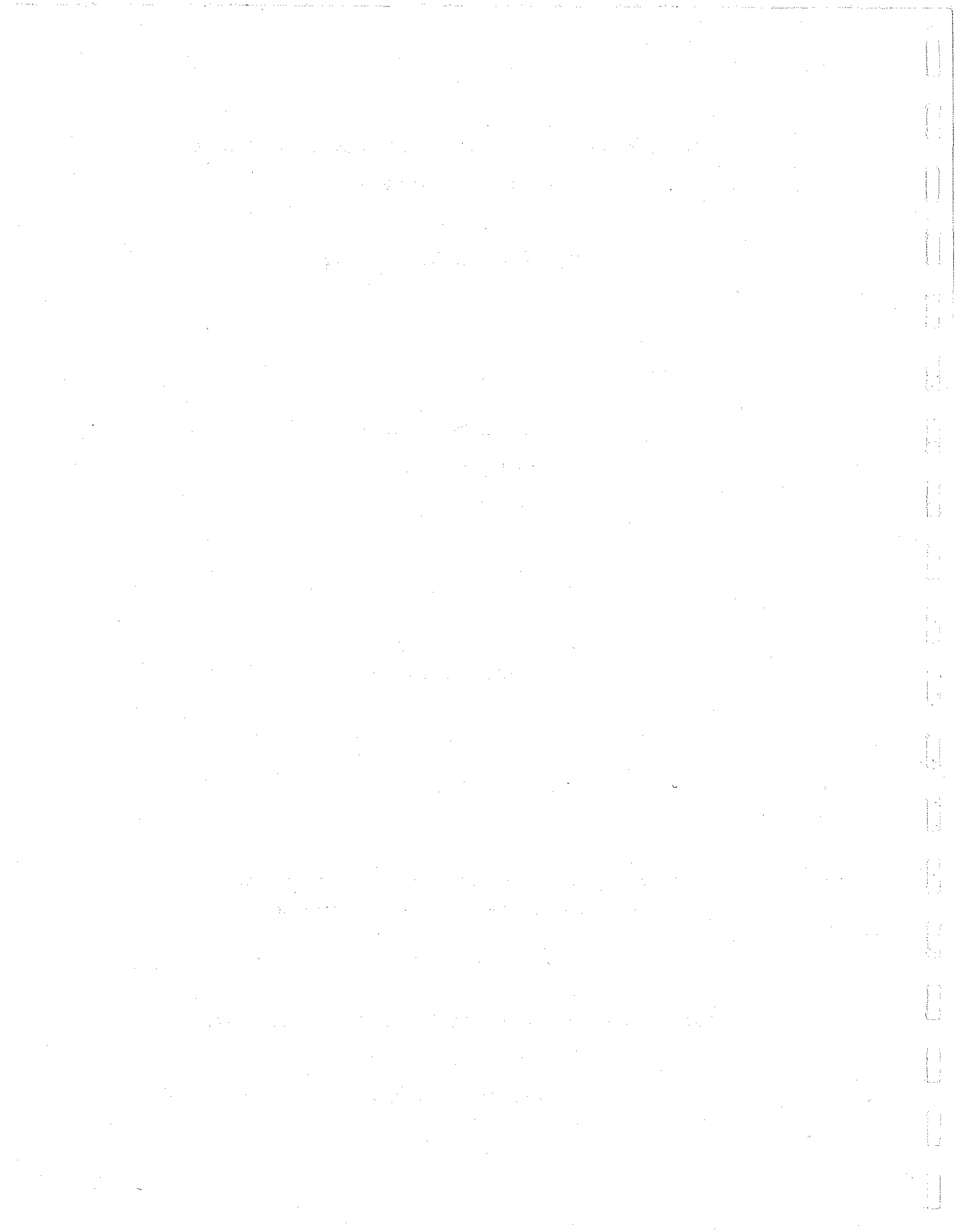
PSYCHOLOGY SERIES

Reproduction in Whole or in Part is Permitted for  
any Purpose of the United States Government

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA



Some Counting Models for First-grade Performance Data on  
Simple Addition Facts<sup>1</sup>

Patrick Suppes and Guy Groen

Introduction. To mathematicians or educators who have not thought about the matter very much, it usually comes as a surprise, and occasionally even as a shock, to find out how little we yet know about the learning of mathematics. It is not uncommon to hear mathematicians say that because mathematics is a systematic subject with an inherent order imposed on the development of topics, it should be relatively straightforward to give quite an adequate account of how students learn mathematics. Because students do learn mathematics and because many of the mathematicians who make this sort of statement have themselves been successful teachers, it is not always evident what is the best way to bring out the gross inadequacies in our present knowledge of mathematics learning.

Perhaps the most effective way, -- at least we have found that it sometimes works --, is to rely heavily on computer analogies. First challenge: if you understand so well how mathematics is learned, please program my computer to learn it. It does not take much discussion to bring out the difficulties of this task, and one can then move on to a second challenge: predict the points at which students will have learning difficulties and make explicit the principles used to make the predictions. The requirement of explicitness is needed to make the

challenge a scientific and theoretical one that cannot be answered by the non-verbalized, intuitive experience of a good teacher. While one's opponent is struggling with this second challenge, a third challenge on performance data can be put ready at hand: predict systematic variations in performance data on mathematical concepts and skills already taught, and again make the principles of prediction explicit. The view that mathematical colleagues may have difficulty giving a serious constructive response to these three challenges is not meant as a criticism of their scientific prowess. The only criticism implied is of the opinion that we already know how to meet these three challenges in any serious way.

The present article is meant to be a small step toward a positive response to the third challenge. From a mathematical standpoint the performance task we have selected is ridiculously simple, that of handling correctly the simple addition facts, with the sums being no greater than 5. From a psychological standpoint, however, this task is not as simple as most of those that lie at the heart of the classical experiments in learning theory. Moreover, attempts to develop mathematically well-defined performance models for even this simple task do not seem to exist in the literature.

We reserve more detailed comments until after we have presented in the next section goodness-of-fit results for five closely related models. A broader conceptual framework for the viewpoint expressed here is to be found in Suppes (1965).

Experimental test of five models. The results we will discuss are from an experiment in which a group of first-grade children in the first half of the school year were asked to solve a set of simple addition problems. Each problem was of the form

$$m + n = \underline{\quad}$$

where  $m + n \leq 5$ . The line was colored red while the rest of the problem was printed in black. The task of each child was to provide the missing number. Thirty subjects were used, randomly selected from two different homerooms. Each subject was run individually. Subjects were seated in front of a panel with six buttons marked 0, 1, 2, 3, 4, 5. They were then shown a sample problem projected on a screen in front of them. They were told that the red line meant that a number was missing and were instructed to push the correct button for the missing number. When a subject had responded, he was shown a new slide with the correct answer (printed in red) replacing the red line. Each child was then presented with a sequence of twenty-one problems consisting of all possible combinations of integers  $m$  and  $n$  subject to the constraints

$$m + n \leq 5$$

$$m \geq 0$$

$$n \geq 0$$

These problems were presented in a random order, the same sequence being used for each child. After each presentation of a problem, the child made a response and was shown the correct answer. Both the actual response and the response latency (i.e., response time) were recorded. This procedure was repeated for two more days. However,

on the last two days, no preliminary instructions were given, and the child was asked to respond as quickly as possible. The order of presentation of items was different on each of the three days.

In this discussion, we will concentrate on the data obtained on the third day. By then, it can be assumed that the children had become fully familiar with the experimental situation. The initial problem we proposed to consider was whether it is possible to formulate a simple model that will account in an approximate fashion for the childrens' responses.

Unfortunately, the error rate was too low for any systematic analysis to be based on this aspect of the response data. Although at least one subject made an error on each problem, seven subjects (out of thirty) made errors on  $1 + 3 = \underline{\quad}$  and  $1 + 2 = \underline{\quad}$ , five subjects made errors on  $4 + 1 = \underline{\quad}$ ,  $3 + 2 = \underline{\quad}$  and  $1 + 1 = \underline{\quad}$ . On most other problems one or two subjects made an error.

As a result of these low error rates, it seemed more promising to consider the response latencies. The most reasonable basic assumption to make is that the variations in response latencies between problems are the reflection of some kind of counting process that the child is using. For a problem of the form  $m + n = \underline{\quad}$ , it is possible to distinguish between five different kinds of counting processes. In order to make this distinction, it is convenient to consider a counter on which two operations are possible: setting the value of the counter to a certain value (while clearing the previous value) and adding a number to the current value of the counter. The addition operation is performed by successively incrementing the counter by one until the value

has been added on. The operation of this counter is illustrated in Fig. 1.

-----  
Insert Figure 1 about here  
-----

Using this counter, an addition problem of the form  $m + n = \underline{\quad}$  can be solved in the following ways:

1. The counter is initially set to 0,  $m$  is added and then  $n$ .
2. The counter is set to  $m$  (i.e., the left-most number) and  $n$  is then added.
3. The counter is set to  $n$ , and  $m$  is then added.
4. The counter is set to the minimum of  $m$  and  $n$ . The maximum is then added.
5. The counter is set to the maximum of  $m$  and  $n$ . The minimum is then added.

The setting operation is assumed to take a constant time independent of the value to which it is set. The addition time, on the other hand, is proportional to the number of times the counter must be incremented. Suppose a counter takes time  $\alpha$  to be set and time  $\beta$  to be incremented by 1. If a counter is to be set to a certain value and then incremented  $x$  times by 1 (which is equivalent to having  $x$  added to it) the total time  $T$  taken by the counter to perform these operations is

$$T = \alpha + \beta x . \tag{1}$$

Thus equation (1) gives the time taken to perform an addition problem of the form  $m + n = \underline{\quad}$ . It will give differential predictions depending

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000



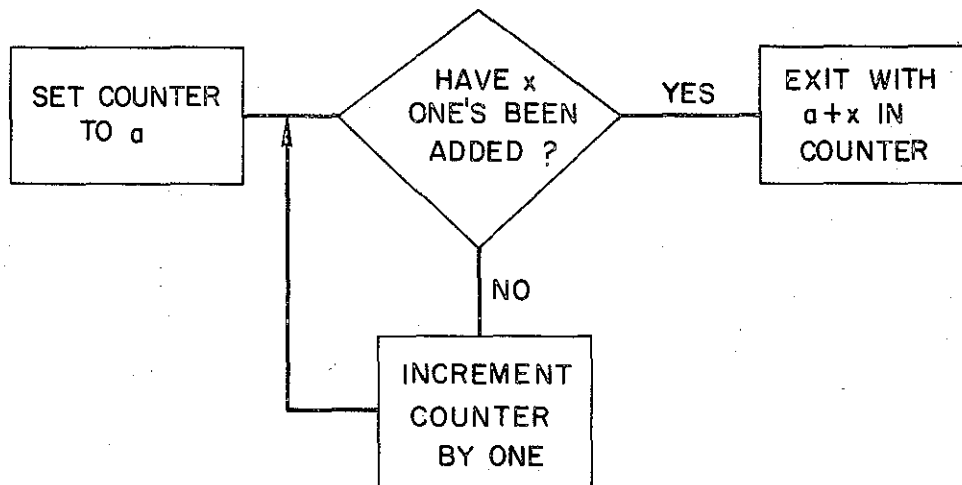


Fig. 1. Example of a device which sets a counter to a and adds x to a.

on the type of solution because, corresponding to the classification of solution types we have just proposed,  $x$  is determined as follows:

Type 1.  $x = m + n$

Type 2.  $x = n$

Type 3.  $x = m$

Type 4.  $x = \max(m, n)$

Type 5.  $x = \min(m, n)$ .

If we wish to apply this model to the latencies of our experimental subjects it cannot be assumed that the values of  $\alpha$  and  $\beta$  are constant. Rather, it is correct to assume that  $\alpha$  and  $\beta$  are random variables with two different distributions. However, we can eliminate this problem by taking the mean latency over all subjects. We then have, for a particular problem  $i$ ,

$$E(T_i) = E(\alpha) + x_i E(\beta) \quad (2)$$

where  $x$  is computed according to the rules given above. For equation (2) to hold, it is necessary, of course, that  $x$  be constant for all subjects on a given problem. In other words, it is necessary to assume that all subjects use the same type of solution. If this assumption is incorrect, then the goodness of fit of observed to predicted data will be affected.

In order to evaluate the goodness of fit of these five models, it is necessary to estimate the expected values  $E(\alpha)$  and  $E(\beta)$ . These estimates will be denoted by  $\hat{\alpha}$  and  $\hat{\beta}$ . For each problem, it is possible to compute a value of  $x$  under each of the five assumptions. Since equation (2) is linear,  $\hat{\alpha}$  and  $\hat{\beta}$  can be computed for each model

by means of a simple regression analysis using  $x_i$  as the independent variable and the observed average success latency on each problem as the dependent variable with the index  $i$  ranging over all twenty-one problems. It is necessary to use the success latency rather than the overall latency for the dependent variable because it is reasonable only to assume that the equation (2) holds for correct solutions.

An analysis of this type was performed on the data obtained on the third day of the experiment. Two problems ( $3 + 0 = \underline{\quad}$  and  $2 + 3 = \underline{\quad}$ ) were omitted from the analysis. On both these problems, many individual response latencies were excessively high. The former was always the first problem to be presented. The high latencies on the latter can also be accounted for on the basis of sequential ordering effects. From the data obtained from the remaining nineteen problems,  $\hat{\alpha}$  and  $\hat{\beta}$  were evaluated for each of the five models, and two indexes of goodness of fit were computed. The first was the mean squared deviation between predicted and observed values:

$$s^2 = \frac{1}{17} \sum_{i=1}^{19} (T_i - \hat{\alpha} - \hat{\beta}x_i)^2$$

where  $T_i$  denotes the observed success latency for problem  $i$ . Also computed was the ratio of  $\hat{\beta}$  to the standard error of  $\hat{\beta}$ . If  $T_i$  is normally distributed then this has a  $t$  distribution with  $n - 2$  degrees of freedom. While it is not entirely clear whether or not the assumptions of the test are satisfied in the present experiment, its application does provide a rough index of whether or not the fit is satisfactory. The values of  $\hat{\alpha}$ ,  $\hat{\beta}$  and  $s^2$  resulting from the analyses of the various models are shown in Table 1. It is clear

-----  
Insert Table 1 about here  
-----

from these results that Model 1 and Model 5 provide the best fits. The second goodness-of-fit computation results in levels of significance beyond .05 for all models but Model 1 and Model 5. The predicted success latencies obtained on the basis of these two models, together with the corresponding observed mean success latencies are shown in Table 2. It is clear from Table 2 that, especially in Model 5,

-----  
Insert Table 2 about here  
-----

each value of  $x$  involves a number of data points. As a result, a clearer notion of the fit can be obtained by comparing the predicted latency with the observed latency averaged over all problems that contribute to a given value of  $x$ . This is done in Fig. 2. Although

-----  
Insert Figure 2 about here  
-----

there are less values of  $x$  in Model 5, the better fit cannot be ascribed to this circumstance, since the value of  $s^2$  is lower for Model 5, despite the fact that it is computed on the basis of deviations between predicted and observed latencies for individual problems.

While it can safely be concluded that Model 5 fits better than Model 1, this result can only be considered to be a first step. There

Table 1. Regression estimates for the different solution types  
 ( $\hat{\alpha}$  and  $\hat{\beta}$  are measured in seconds).

	Model	$\alpha$	$\beta$	$s^2$
1.	$x = m + n$	2.96	.216	.369
2.	$x = n$	3.50	.098	.465
3.	$x = m$	3.48	.119	.404
4.	$x = \max(m, n)$	3.43	.092	.471
5.	$x = \min(m, n)$	3.26	.710	.233

Table 2

Model 1:  $x = m + n$ Model 5:  $x = \min(m, n)$ 

Problem	x	Mean Success Latency (in seconds)		Problem	x	Mean Success Latency (in seconds)	
		Pred.	Obs.			Pred.	Obs.
0 + 0	0	2.96	2.98	0 + 0	0	3.26	2.98
0 + 1	1	3.18	3.36	0 + 1	0	3.26	3.36
1 + 0	1	3.18	3.27	1 + 0	0	3.26	3.27
0 + 2	2	3.40	3.57	0 + 2	0	3.26	3.57
1 + 1	2	3.40	2.67	2 + 0	0	3.26	2.88
2 + 0	2	3.40	2.88	0 + 3	0	3.26	3.45
0 + 3	3	3.61	3.45	0 + 4	0	3.26	3.48
1 + 2	3	3.61	4.20	4 + 0	0	3.26	3.40
2 + 1	3	3.61	4.28	0 + 5	0	3.26	2.85
0 + 4	4	3.83	3.48	5 + 0	0	3.26	3.03
1 + 3	4	3.83	4.18	1 + 1	1	3.97	2.67
2 + 2	4	3.83	3.90	1 + 2	1	3.97	4.20
3 + 1	4	3.83	4.04	2 + 1	1	3.97	4.28
4 + 0	4	3.83	3.40	1 + 3	1	3.97	4.18
0 + 5	5	4.05	2.85	3 + 1	1	3.97	4.04
1 + 4	5	4.05	4.49	1 + 4	1	3.97	4.49
3 + 3	5	4.05	5.15	4 + 1	1	3.97	4.53
4 + 1	5	4.05	4.53	2 + 2	2	4.68	3.90
5 + 0	5	4.05	3.03	3 + 2	2	4.68	5.15

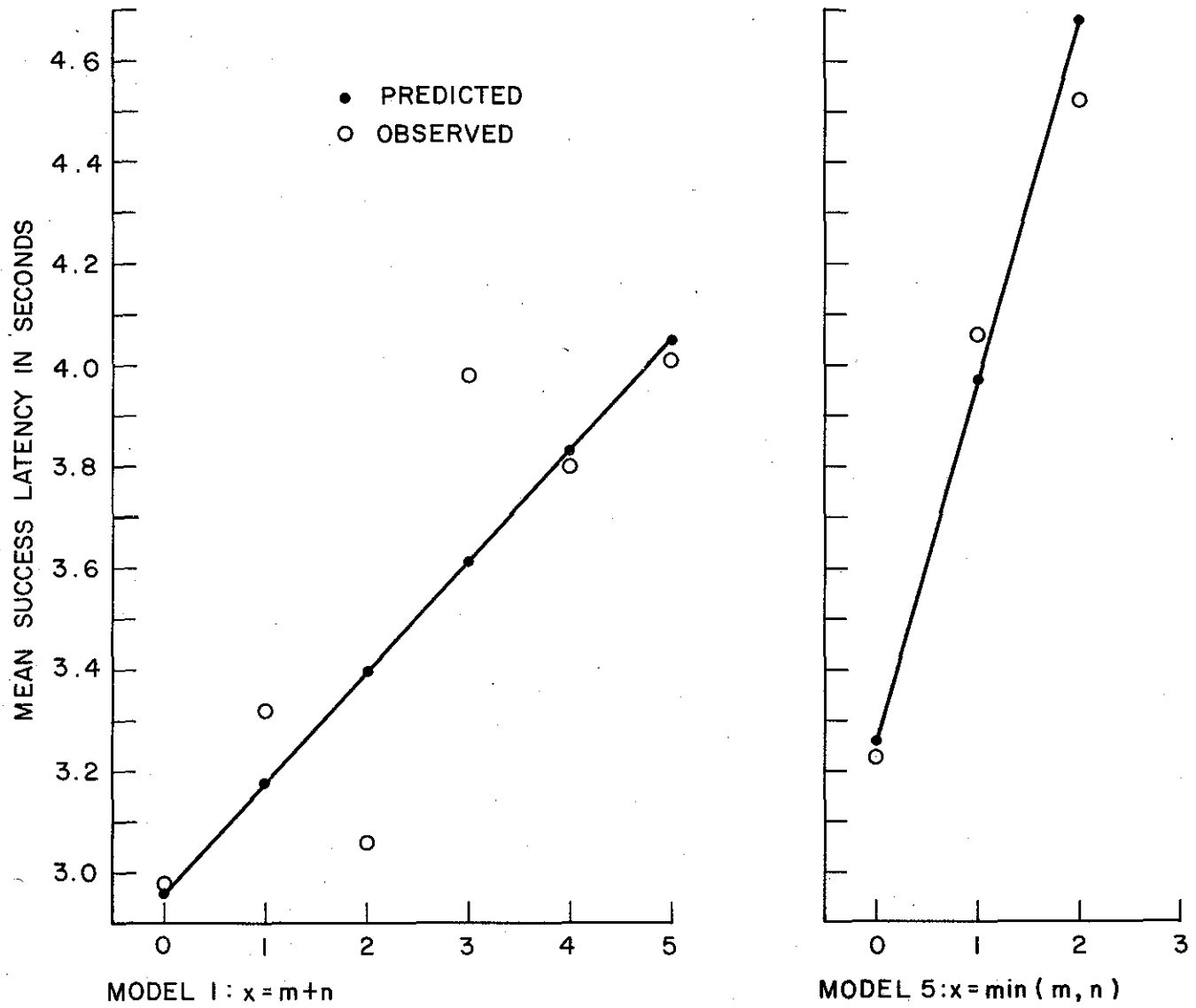
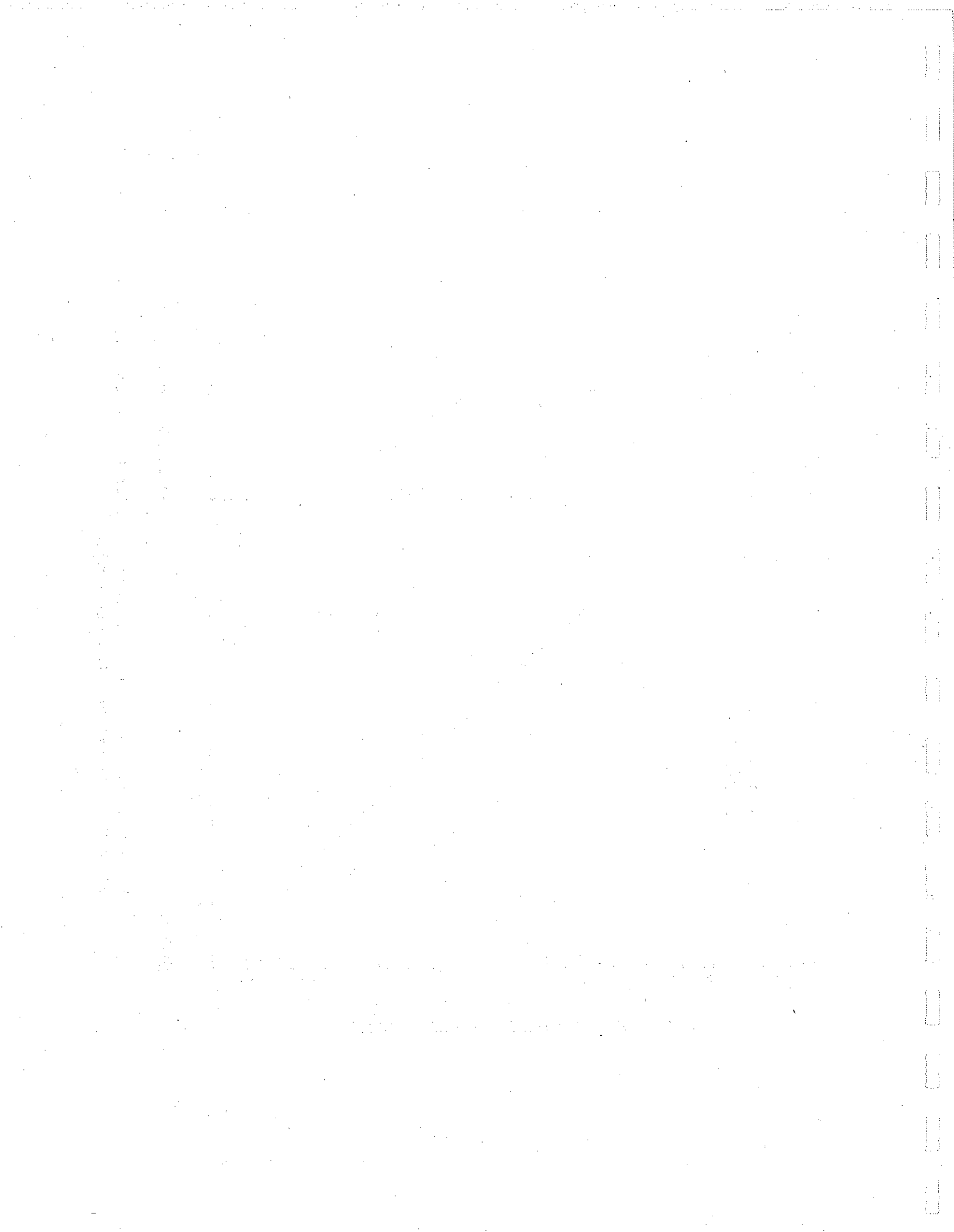


Fig. 2. Predicted and observed mean success latencies plotted as a function of  $x$  for the two best-fitting models.





is no guarantee that no other model exists that would fit the data in a more satisfactory fashion. Moreover, it cannot be inferred that the good fit of Model 5 implies that subjects tend to add two numbers according to the mechanism suggested by the model. For this model,  $x$  ranges from 0 to 2. It is only when  $x = 2$  that neither a 0 nor a 1 appears in the problem. Hence the data might be accounted for by a model which assumes specific algorithms for solving the problem when a 0 or 1 appear in it rather than the general algorithms used by the models we have proposed in this paper. Subsequent research that deals with these matters is now planned.

Some concluding remarks. It would be good if we could report that the algorithm represented by Model 5 was the one explicitly taught the children by their teachers. This does not seem to have been the case. At the present time most first-grade teachers do not teach their students an explicit counting algorithm for handling the simple addition facts ordinarily taught in the first grade. As would be expected there is usually some mention, and often even a fair amount of discussion, of counting and its relation to the first introduction of addition. But -- and this is the important point -- an explicit algorithm is not developed and taught as is done later for addition of multi-digit numbers.

The results of the present paper suggest that more attention might profitably be devoted to these first algorithms, and that the algorithm of Model 5, which seems more sophisticated than that of Model 1, might well receive more explicit emphasis in the teaching of first-grade arithmetic.

It has not been our intention in this short paper to present any definitive research, but only to illustrate how even so simple a thing as learning the addition facts presents an interesting challenge to learning theorists and affords already an opportunity to test some alternative mathematical models, each of which rests on a clear intuition of how a simple addition problem may be solved. The central idea of a counting model seems so natural that it seems difficult to think of other possible approaches, but this is not really the case -- for example, a table-look-up model with parameters appropriately introduced for scanning the table can be formulated in such a way that it is identical in all behavioral predictions with Model 5. Moreover, simple counting ideas are not sufficient to account for all the significant variations in the observed data of Table 2, and as a larger body of data is accumulated, more complex and subtle ideas will be needed in constructing an adequate model of the observed phenomena. On the other hand, it seems to us that the learning of elementary mathematics affords a natural testing ground for mathematical models of learning or performance, and there is some reason to hope that in a first approximation at least models of a reasonable degree of simplicity will suffice.

It should be apparent that as such models are developed and the range and depth of their success increased they will have increasing significance in suggesting and guiding curriculum modifications, particularly in the direction of concern for the fundamental problem of finding out how students can on the average best learn mathematical concepts and skills.

## Reference

P. Suppes, The Psychological Foundation of Mathematics, Technical Report No. 80, Psychology Series, Institute for Mathematical Studies in the Social Sciences, Stanford University, October, 1965. 33 pp.

## Footnote

<sup>1</sup>This research was supported by the United States Office of Education, The National Science Foundation and the Carnegie Corporation of New York. We would like to thank Mrs. Sue Matheson for running the experiments in the school.

