The Journal of General Psychology, 1950, **42**, 279-297.

NEW EXPERIMENTAL ATTEMPTS AT PREVENTING MECHANIZATION IN PROBLEM SOLVING*

Department of Psychology, McGill University

ABRAHAM S. LUCHINS AND EDITH H. LUCHINS

A. INTRODUCTION

More than a decade ago we began a series of extensive investigations of the possible deleterious effects of habituated behavior (4, 5). We have been concerned with the blinding effects of habit, with what happens when a habit "ceases to be a tool discriminantly applied but becomes a procrustean bed to which the situation must conform; when, in a word, instead of the individual mastering the habit, the habit masters the individual" (4, p. 93). Specifically, our studies have dealt with the tendency to impute a procedure, repeated in a series of similar tasks, to subsequent problems which possess more direct solutions. This tendency toward mechanization has been taken as evidence of an Einstellung (or a special kind of mental set). Various experimental factors introduced to prevent an Einstellung from developing or to weaken it after it occurred, have on the whole been quite ineffective. The present paper deals with some new experimental attempts to prevent or overcome this Einstellung tendency.

In order to facilitate understanding of the present experimental variations, we shall first briefly describe the original basic experiment upon which they are founded (4, pp. 1-17; 3).

B. BASIC EXPERIMENT

The experiment involves volume-measuring problems, in each of which a certain number of empty containers and a supply of fluid are considered to be furnished. By means of these, with the use of pencil and paper, *S* is to figure out how to obtain a stipulated volume of fluid. The problems, presented at intervals of about 2½ minutes, are given in Table 1.

The first task is for illustrative purposes. If the jars, in the order written, are labeled with the letters *a, b, c,* respectively, then Problems 2 through 6 are solvable by the formula $b - a - 2c$; e.g., $127 - 21 - 2 \times 3$ equals $127 - 27$ equals 100 gives the solution of Problem 2. The next two tasks are solvable by this formula and by other simpler ones; $a - c$ and $a + c$,

*Received in the Editorial Office on November 8, 1948.

TABLE 1
THE TASKS

Problem	Containers given (capacity in quarts) a	b	c	To get
(1)	29	3		20 quarts
(2)	21	127	3	100 "
(3)	14	163	25	99 "
(4)	18	43	10	5 "
(5)	9	42	6	21 "
(6)	20	59	4	31 "
(7)	23	49	3	20 "
(8)	15	39	3	18 "
(9)	28	76	3	25 "
(10)	18	48	4	22 "
(11)	14	36	8	6 "

respectively. The ninth problem does not fit the formula $b - a - 2c$ but is solved by $a - c$. Problems 10 and 11 are similar to the seventh and eighth.

Problems 2 through 6 are the Einstellung problems, and Problems 7 through 11, which test their effect, are the test problems or the criticals. The method represented by the formula $b - a - 2c$ is called the Einstellung method: the procedures represented by the formulas $a - c$ and $a + c$ are named the direct methods.

Measures of the amount of mechanization or set, the amount of Einstellung effect, are given by the percentage of Einstellung solutions of the first two criticals (considered as a unit and called C_1C_2); the percentage of Einstellung solutions of the last two criticals (considered as a unit and called C_3C_4); and the percentage failure of Problem 9. A measure of the amount of recovery from mechanization is given by the increase in direct solutions in C_3C_4 as compared with C_1C_2.

This basic experiment and its variations have been administered by the author to over 9,000 *S*'s. In computing the results we considered the responses of only those *S*'s who solved at least the last two Einstellung problems (the two problems immediately preceding the criticals) by the Einstellung method,[1] since we were interested in the effect that such solutions had upon subsequent solutions of the criticals. Most of these *S*'s showed considerable Einstellung effect. Recovery from mechanization was in general not large for adult groups, and was negligible in most elementary school groups.

[1]Solutions of these problems which utilize procedures other than the Einstellung method are possible, and were obtained in a few cases.

C. Limiting the Amount of Fluid

The Einstellung procedure is more wasteful of fluid than are the direct methods of solution, in the sense that it requires the discarding of a larger volume of fluid. For example, in the second critical—given a 15-quart jar, a 39-quart jar, and a 3-quart jar, obtain 18 quarts of fluid—the Einstellung method necessitates beginning with 39 quarts of which only 18 are actually utilized while 21 quarts are discarded, but the direct method uses only the 18 quarts demanded by the solution.

In an attempt to further the use of the direct procedures in the criticals, we decided to emphasize the need for finding methods of solution which were not wasteful of fluid. Since the search for economical methods would be realistic only if the amount of fluid available was limited, it was decided to alter the basic setup, in which the supply of fluid was well nigh unlimited, and to set the volume available at 539 quarts. This volume suffices to solve all the 10 problems after the introductory one, if *S* employs the Einstellung procedure in the Einstellung problems and the direct methods in all criticals. As was the case when the original setup was administered to elementary school children, the fluid was said to be milk.

After the introductory remarks and the first problem were presented as in the basic experiment, *S*'s of the present variation were told:

> You will receive 10 more problems. In each, your task will be to measure the amount of milk asked for by a customer, using as measures any or all of the empty containers he gives to you. There is a large tank from which you may take out the milk but, for sanitary reasons, once you put milk in a certain customer's containers you cannot pour it back into the tank or use it for anyone else; you will have to throw away any milk you do not actually give to him, once it is in his containers. The tank contains 539 quarts of milk which will satisfy the orders of all 10 customers, if you are careful to use methods which waste as little milk as possible.
>
> When you finish a problem, figure out how much milk there is left in the tank. Write down the amount in the space allowed for the next problem so that you will be able to see whether there is enough milk left with which to solve that problem.

After these instructions, the first Einstellung problem was presented and 2½ minutes allowed for its solution. Then, as in the basic experiment, the two methods of solving this task[2] were illustrated. In addition, in each case the class was shown how to determine and record the amount of milk

[2]$127 - 21 - 2 \times 3 = 127 - 27 = 100$ and $127 - 9 \times 3 = 127 - 27 = 100$.

remaining in the tank. The subsequent problems were then presented as in the basic experiment.

It was hoped that the knowledge that he must not needlessly waste the fluid would produce a persistent drive on *S*'s part to search for economical, direct solutions, and that in keeping track of available fluid, *S* would be reminded of the need for conserving the dwindling supply by utilizing economical procedures. An *S* who kept account of available fluid and used the Einstellung method in the Einstellung problems and the first two criticals, would find that he was left with only 17 quarts in the tank, insufficient to solve the remaining tasks in the Einstellung manner, and only enough to solve the last problem in the direct way. The repeated failure to use the Einstellung method might help disrupt the mental set and cause recovery in C_4.

Finally, while an *S* in the basic experiment might be satisfied with solving all the problems but one or a few, *S*'s in the present variation might feel a deeper sense of failure if unable to do some of the criticals, since the assigned task was to solve each problem so that sufficient fluid would remain for the subsequent ones. For these reasons, it was believed that the present variation might be successful in operating against Einstellung effects.

1. *Experiment I*

The described setup was administered to four classes in the sixth year of a New York City public elementary school. Excluded from consideration are the responses of 20 *S*'s who did not solve the last two Einstellung problems. Of the remaining 139 *S*'s, only 33 obeyed instructions and consistently computed and recorded the volume of available fluid; with minor exceptions the computations involved were correct. These 33 showed only 35, 15, and 5 per cent Einstellung solutions[3] of C_1C_2, #9, C_3C_4, respectively, while the 106 who failed to keep the record throughout (most of them stopped after the first two or three problems) contributed 81, 76, and 79 per cent Einstellung solutions of these problems. Differences between the two groups of *S*'s are 46, 61, and 74 per cent, with *CR*'s of 7.2, 8.1, and 19.1, respectively, indicating statistical reliability. The 30 per cent recovery shown by those who kept track of available fluid contrasts sharply with the 2 per cent made by those who failed to keep the record.

Considering the group as a whole, we find that this experimental variation was not very successful in operating against Einstellung effects, since the 139 *S*'s gave 71, 62, and 62 per cent Einstellung solutions of C_1C_2, #9, C_3C_4, respectively— results which fall within the range of responses made by similar sixth year groups which had participated in the basic experiment.

Comments of *S*'s after the experiment furnished varied explanations for the failure of so many consistently to keep account of available fluid. Some claimed that they required the allotted 2½ minutes for a problem, and did not have time to compute the amount of remaining milk. Others said that they became so engrossed in working on the problems that they forgot the instructions. A few admitted that while they remembered the directions, and even complied with them initially, they did not trouble to keep the record thereafter, since it seemed superfluous and a waste of time; as long as they were able to obtain the correct answers to the problems, they saw no point in determining the contents of the tank.

While those who kept track of the remaining milk showed comparatively little Einstellung, one might ask why they showed any. It was our impression, based on discussions with the *S*'s, that some of those who did keep the record of the fluid did so in quite mechanical a fashion, without relating the computations to their methods of solution of the problems. They often paid scant attention to the amount of available milk before beginning the next problems, or, in a few cases, ignored this figure if it indicated to them that there was insufficient milk to permit the utilization of their method, reasoning that it was better to "solve" the remaining problems under these conditions than not to solve them at all, or that perhaps they had erred in their record-keeping.

2. *Experiment II*

Since so many *S*'s did not keep track of the fluid in the previous experiment, it was decided to vary that setup by reminding the class of this instruction *after every problem*. In addition, those whose records after the second critical revealed that they did not have sufficient fluid with which to solve the three remaining tasks, were given the opportunity to return to the first two criticals and to search for less wasteful solutions. When the ninth problem had been on the blackboard for 2½ minutes, we said:

> Please stop working and listen carefully. Before beginning this last example you should have had 64 quarts of milk in the tank. If you did not, then take the extra sheet of paper which was given to all of you before, and try to do over the last few problems in order to find ways which will waste less of the milk. After that, do the ninth, tenth, and eleventh examples, which will be left on the blackboard long enough so that you will have time to work them. If your record

[3]When it simplifies the presentation of data, we shall refer to failures of Problem 9 as Einstellung solutions.

just before the ninth problem showed that the tank contained 64 quarts, then you need not do the previous examples over but may go on to the tenth and eleventh.

Problems 9, C_3, and C_4, written on the blackboard at the usual 2½-minute intervals, were allowed to remain there for 15 minutes.

This setup was utilized in two sixth-year classes and resulted in each of the 60 *S*'s keeping a consistent, fairly accurate account of available fluid. Nonetheless, they showed 80 and 73 per cent Einstellung solutions of C_1C_2 and #9, prior to the interruption after the latter task. These results, similar to the range of responses made by sixth-year classes in the basic experiment, hint that the record-keeping *per se* did not induce the use of direct solutions, that there was not, in general, a causal relationship between keeping the record and searching for economical methods. Discussions with the *S*'s revealed that most of them computed the volume of available milk in a highly mechanical fashion, because they were reminded to do so, but that the computations had little significance to them; previous experiments (4, pp. 60-63) also have indicated that a reiterated instruction of this sort tends to become meaningless to the *S*'s. In the present experiment many admitted that they actually did not consult the computations before solving the subsequent problem or that they ignored it if it did not support the use of the Einstellung method.

Subsequent to the interruption following the ninth problem, after which *S*'s had the opportunity to do some of the problems again, the 60 *S*'s gave only 17, 13, and 13 per cent Einstellung solutions of C_1C_2, No. 9, and C_3C_1, respectively. Differences in Einstellung solutions were 63 per cent in C_1C_2, with a *CR* of 12.6, and 60 per cent in Problem 9, with a *CR* of 8.4, both statistically reliable. These differences in results testify to the efficacy of centering *S*'s, after the extinction task, on the possible lack of fluid, reminding them that wasteful methods might have produced this lack, and giving them another chance to solve some of the problems. The success of the interruption after the ninth task further supports the assumption that an individual tends to persist in the use of a stereotyped mode of response until some dramatic occurrence causes him to start searching for new methods (cf. 1, p. 342).

3. *Experiment III*

In the previous experiment considerable recovery occurred when *S*'s were given the opportunity to do some of the criticals again. We were interested in matching a similar opportunity against "speed conditions." Accordingly, as in previous speed experiments (4, pp. 53-56) *S*'s were first told that they were being tested to see how quickly they could solve the 11 problems (all of which had already been written on the blackboard), that they must work as quickly as they could, and bring their papers to *E* as soon as they finished so that he might record the time of completion. At frequent intervals, they were told that they must hurry, that younger children worked faster, etc. In other respects the initial instructions were similar to those of Experiment I, including the restriction against returning milk to the tank and the direction to compute and record the amount of remaining fluid.

This setup was administered to two sixth-year classes. It was found that only four out of 62 *S*'s kept a consistent record of the remaining fluid. These four showed no Einstellung effect, while the remaining 58 *S*'s gave 100 per cent Einstellung solutions of all criticals. As each of the latter brought over his paper, *E* told him that he had failed to keep tab of the amount of milk remaining in the tank and that because of the wasteful methods he used, he actually had not had enough fluid to do even the ninth problem. He was given another sheet of paper, told to do the problems again, to keep a record, and to pay particular heed to the last five problems.

On their second papers, only 16 of the 58 showed recovery, so that the latter now gave 73 per cent Einstellung solutions of each critical—which, while it differed significantly from the previous 100 per cent, was as large as the Einstellung effects yielded by some sixth-year groups which received the basic setup. It is noteworthy that the 16 *S*'s employed the direct solution in every critical while the others never utilized this procedure; they repeated the Einstellung method throughout, even using it in Problem 9 where it did not yield the desired volume of water and did not at all solve the problem. It would seem from the results that a dramatic occurrence which starts one searching for new methods does not necessarily reduce Einstellung effects. A possible explanation for this seems to lie in the subjects' reactions to the situation.

Most of the *S*'s who had to do their work over later told us that they were disturbed and humiliated by *E*'s announcement that they had not solved the problems correctly and must do them again, since it meant that in spite of all their hurrying they would now be among the last to finish and might even fail the test. They said that when they attempted to do the problems over, they could not relax and think clearly; they could think of nothing but their failure and the need for speed. Some of these *S*'s tended to submit their second papers within a few minutes, often running all the distance to *E* in order to save a few seconds; many were clearly under the influence of the "speed atmosphere," the effects of which were also seen in strained faces, broken pencil points, and occasional tears.

4. *Discussion and Implications*

1. It is our impression (corroborated by *S*'s comments) that the initial instructions focused some *S*'s on the need for conserving the amount of milk through the utilization of economic methods, but that their inability to find a more economical procedure than the Einstellung method in the first few problems led them to abandon their search for other methods as a waste of time, or led them to conclude that the Einstellung procedure was the least wasteful or, in fact, the only method of dealing with these problems. Finally, it could have been that even though *S* was initially concerned with discovering more economical methods, the repeated use of the Einstellung method in the Einstellung problems developed in him a mechanical tendency to repeat it, a tendency which swept in its wake any consideration of other methods. As opposed to these *S*'s, were those who paid no heed at all to the need for discovering economical procedures but who from the start fell victim to the Einstellung tendency. In other words, in this experiment as in previous ones, distinctions must be drawn among Einstellung effects (*a*) which are the result of a conscious generalization, (*b*) which result from sheer mechanical perseveration, and (*c*) which initially arise from a generalization but change into mechanical repetition. This again raises the crucial problem of determining the precise nature or natures of the Einstellung phenomena (cf. 4, pp. 87-89).

2. We wonder whether the lack of success of the variation employed in Experiment I is not in some ways an outcome of the fact that we used arithmetical problems, and that the *S*'s, elementary school children, had developed certain attitudes and habits toward arithmetic as a result of their schooling. They were accustomed to the use of isolated drill in arithmetic, wherein in order to "learn" a method or a formula they practiced it in a series of similar problems—a situation quite similar to our experimental setup. They were accustomed to being taught a method and then practicing it; to have to discover procedures was not only quite foreign to them in arithmetic but also in most school subjects. It seems to us that the methods of teaching to which they had been subjected tended to develop, not adaptive responses, but fixations, so that a child might know methods and formulas and yet not know where to apply them or how to determine what method best suited a particular problem. Our schools may be concentrating so much on having the child master the habits, that the habits are mastering the child.

Perhaps what we found in our experiments was an outcome of such schooling (cf. 4, pp. 90-93).

3. The most effective technique was that utilized in Experiment II, in which *S*'s were reminded, after the ninth task, of the amount of fluid there should have been in the tank prior to beginning this task, and given the opportunity to do over some of the previous problems if they did not have at least this amount. Under speed conditions, however, in Experiment III, the opportunity to do the problems over did not produce substantial recovery. We believe that the difference in results may be attributable to differences in social atmospheres prevailing under the two experimental conditions.

Most public school pupils who participated in the basic experiment and its numerous variations tended to have a superficial relation to the problems, perhaps due to school-fostered attitudes towards problem solving, grades, and tests; they were not interested in the problems *per se,* but only in how high a grade they would make, whether the teacher or principal would see their papers, and if their scores would affect their report card marks, etc. This superficial relationship was intensified under speed conditions, with its emphasis on haste and competition. After hurrying to finish and to bring his papers to *E*, the latter's announcement that he had not solved the problems correctly and must do them over was an emotionally upsetting, frustrating experience for *S*. He returned to his seat in no frame of mind calmly to examine the problems and survey methods of solution; the humiliation of having failed the "test" in his first attempt, the knowledge that he must hurry in doing the problems over, and the fear that he would now be among the last to finish, were not conducive to orderly thinking, but tended to narrow *S*'s mental horizon and blind him to new solutions, so that about all he could do was to repeat the oft-repeated Einstellung method.

Thus, the problem solving situation under speed conditions was one in which the main motivating drives for solution came from external pressures on *S* and from ego-motivated forces, rather than from the problems themselves. Indeed, the problems played only a peripheral rôle in *S*'s cognitive grasp of the situation, so that their structural requirements were not given a chance to guide him. While this was true to some extent even in Experiment II, it was so to a smaller degree, and the interruption after the ninth task seemed to weaken the superficial attitude toward the tasks for many *S*'s, to cause them to center on the problems and view them anew. *E*'s critique under speed conditions in general only heightened the superficial relationship between *S* and the problems, since it increased the strength of ego-motivated forces and social pressures.

This raises the issue of the rôle of motivation and drive in learning. Firstly, it is necessary to differentiate between motivations which stem primarily from the problems themselves and those which arise from the ego and other pressures extrinsic to the problems. While motivation is generally necessary for problem solving, when the motivation becomes very dominant, particularly if it is of the latter kind, it may force the problem to a peripheral position in the problem solving situation, and cause *S* to lose his docility and variability, to become fixed and rigid in his behavior and insensitive to the requirements of the problem.[4]

We wonder whether our schools with their stress on grades, their test tensions, and competitive atmospheres, are not conducive to an emotional, highly ego-involved approach toward problem-solving, and consequently, whether they are not detrimental to productivity and flexibility in thinking. In a broader sense, are not deleterious effects perhaps exerted upon thinking and emotional balance by the speed emphasized in our industrial systems, our athletic events, and in our society at large?

One further educational implication: The effect of *E*'s critique in the speed experiment, which appeared to frustrate many *S*'s, supports Hilgard's observation that "frustrating agents must be avoided in learning if abnormally rigid habits are to be avoided and teachability not interfered with" (1, p. 307).

4. When we noticed that most *S*'s did not keep a record of available fluid (except in Experiment II), that many who kept the record did not bother to consult it, and that some merrily continued to use the Einstellung procedure even when aware that their record did not permit its use, we could not help thinking that the children might be reflecting the behavior of those adults in our society who do not budget their finances, or who carefully work out a budget and then ignore it, and finally of those who live beyond their means, who spend more than they earn.

It is interesting to speculate whether the general lack of success of this variation is in any way traceable to our nation's notorious lack of conservation of supplies and resources. Perhaps the procedures would be more successful if they were employed with groups which were less wasteful of resources or specifically, with children who hungered for milk, youngsters who are all too plentiful in famine-swept Europe and Asia. Future experimentation may furnish the answer.

[4]Similar results were found in the writer's unpublished studies of the influence of needs and social field conditions on perception.

D. Addition of a Fourth Jar

Essentially, the main difference between the basic setup and the present one was the addition in each problem of a fourth jar not required by the Einstellung method. This change was introduced because we had noticed that *S*'s who participated in the original study often imposed the Einstellung procedure upon a problem without first surveying all the given containers. It appeared that they viewed the problems from the frame of reference of this procedure, since often they immediately selected the largest jar, the center jar, and removed excess fluid from it with the two remaining containers. Maintenance of this frame of reference was strengthened by the original setup since each problem, after the introductory one, consisted of three jars and the Einstellung method demanded just this number. That the tasks in the present variation involved four jars might lessen the likelihood of *S* rapidly imputing this method to them. Even if set to employ the Einstellung procedure, he would first have to examine the four containers in order to select the ones to utilize; in the course of so doing he might become aware of the direct solution or of other solutions made possible by the addition of the fourth jar.

The content of the superfluous container was given as 89, 43, 27, 37, 22, 15, 16, 37, 16, and 4 quarts in Problems 2-11, respectively. Preliminary experimentation revealed that Einstellung effects were as large as in the basic experiment when the additional jar was consistently placed in the same position with relation to the other containers, particularly when it was always first or always last. It was therefore decided to vary the position of this jar which was presented as the 3rd, 2nd, 4th, 3rd, 4th, 2nd, 1st, 3rd, 4th, and 3rd in Problems 2-11, respectively.

1. *Experiment I*

The variation was administered in two New York City public elementary schools to 148 children in four classes on the sixth-year level, classes similar in composition to sixth-year groups which had participated in the basic experiment. The group as a whole (including those who failed to use the Einstellung method in both of the two problems immediately preceding the criticals) gave from 4 to 11 per cent Einstellung solutions and from 80 to 90 per cent direct solutions of the criticals, quite the reverse of results usually shown by similar groups which participated in the basic experiment. Those 28 *S*'s who did employ the Einstellung method in both the two problems prior to the criticals gave 28, 7, and 10 per cent Einstellung solutions of C_1C_2, #9, and C_3C_4, respectively. This was more Einstellung effect than the group as a whole, but significantly less than that shown by a comparative group of 36 *S*'s which received the basic setup; the differences in Einstellung effect in these problems were 44, 54, and 62 per cent with critical ratios of 5.5, 5.7, and 9.4, respectively, indicating statistical reliability. It is interesting to compare the 18 per cent recovery from mechanization made by the 28 *S*'s, with the complete lack of recovery in the comparative group.

Of the 148 *S*'s who received the four-jar setup, 120 failed to employ the Einstellung method in the last two Einstellung problems. There were 24 *S*'s who did solve these two problems, but who used another procedure, frequently a round-about variation of the Einstellung method (e.g., $163 - 25 - 43 - 14 - 25 + 14 + 43 - 14 = 99$), in one or both of them. The remaining 96 *S*'s, constituting 65 per cent of the entire group, failed to solve these two problems, in contrast to the average 12 per cent failure in the sixth-year public elementary school groups of the basic experiment.

In short, the increase of direct solutions of the criticals was offset by an increase in failures and inefficient solutions of the Einstellung problems.

2. *Experiment II*

The four-jar setup was also administered to 125 high school students in three classes on the third and fourth year levels of the High School of Science in New York City. They constituted a select group with regard to intelligence and to aptitude for mathematics or science. The group as a whole (including those who did not employ the Einstellung method in the two problems prior to the criticals) showed from 11 to 28 per cent Einstellung solutions and from 59 to 74 per cent direct solutions of the criticals, about the reverse of results made by high school groups of the basic experiment. Those 71 *S*'s who did employ the Einstellung procedure in these two problems gave 46, 19, and 16 per cent Einstellung solutions of C_1C_2, #9, and C_3C_4, respectively. This was more Einstellung effect than the group as a whole but significantly less than that made by a comparative group of 42 *S*'s of the same school which received the basic setup; differences in Einstellung effect in these problems were 28, 30, and 45 per cent, with *CR*'s of 4.3, 3.3, and 7.4, respectively. The 30 per cent recovery from mechanization shown by the 71 *S*'s compared favorably with the 13 per cent made by the comparative group.

There were 54 *S*'s receiving the four-jar setup who did not use the Einstellung method in the last two Einstellung problems. Of these, 19 solved these problems but used other methods, usually cumbersome variations of the Einstellung procedure, in one or both of them. The remaining 35 *S*'s failed to solve at least one of these problems, in contrast to their almost universal solution in high school groups receiving the basic setup.

Thus, in comparison with the basic experiment, the four-jar variation again yielded more direct solutions of the criticals and (although not to as marked a degree as in the elementary school group) more inefficient solutions and failures of the Einstellung problems.

3. *Discussion and Implications*

In an attempt to understand the nature of solution of the criticals, we questioned the *S*'s after the experiment. The replies of some of those who employed the direct methods in these problems substantiated the hypothesis raised in the introduction, i.e., while they were examining a problem the direct procedure suddenly confronted them, "stared them in the face," even though they had previously used the Einstellung method or variation of it. For some this occurred in the first critical, and for others, after their experience with Problem 9.

Those who failed to employ the direct method in any or all of the criticals, either could offer no explanations for their behavior, simply stated that they did not see the direct solutions or that they had become accustomed to the longer solutions, or said that they had learned in the earlier problems to focus on the largest jar or on a subtractive process.

There are several plausible explanations of the inefficient solutions and failures of the Einstellung tasks, explanations rooted both in theoretical considerations and in some *S*'s comments. To begin with, many *S*'s developed a vague idea of beginning with the largest jar[5] and pouring from it, depending on a "hit and hope" method to obtain the required volume. This resulted frequently in clumsy variations of the Einstellung procedures, involving unnecessary steps which negated each other. Insufficient time to carry out the numerous steps or an error in calculation somewhere along the way contributed to failures, particularly in the elementary school classes.

Moreover, the addition of the fourth jar made the problems look more complicated and actually increased their complexity by increasing the number of possible numerical combinations of the containers. In the basic experiments, *S* was faced with problems which were less complex both in

[5]Preliminary experiments in which this possibility was ruled out show somewhat smaller percentages of Einstellung effects but greater percentages of failures to solve the Einstellung problems.

appearance and in structure, since to use the Einstellung procedure in them required all of the three given containers. In order to employ the Einstellung method in the four-jar variation, *S* had to discard one, only one, and the proper one of the containers. In selecting the jar to be filled initially, and in selecting the jars with which to pour, he might choose the superfluous containers—and err. Nothing in the statement of the problem told how many and which jars should be employed in the solution.

Finally, the basic setup had a familiar ring to the *S*'s because of its similarity to arithmetical isolated drill exercises to which they were well accustomed. To be given problems in which it was necessary to discard one of the givens was an unusual assignment for them. They had been taught to use all the hypotheses given in a problem; indeed, some of the high school students had learned in their geometry classes to check off each hypothesis as they used it, and if any were left unchecked they knew that they had not proceeded correctly. While in natural problem solving situations the selection of facts and hypotheses from the many available ones is an important aspect of the problem solving process, it appeared to be a highly artificial procedure to many of our *S*'s because of the nature of their school training.

This raises the problem of the deleterious effects of drill in mathematics and other school subjects, and the even broader issue of how to teach our students to cope with actual problem situations. An attempt to teach selection of hypotheses underlies the use in certain algebra curricula of problems which have insufficient or superfluous hypotheses. But even the limited value which can arise from the isolated use of such problems in one subject-matter is almost completely negated by the nature of these problems and by the manner in which they are presented. Firstly, they are often taught as optionals, which are squeezed in near the close of the semester if time permits. Secondly, these problems are generally introduced as possessing one too many or one too few hypotheses, and the one in question is often so evidently conspicuous because of its presence or absence that little analysis is required to detect it. Also, these problems are sometimes based on familiar "type-exercises," in which the students previously had much practice, so that anything lacking or superfluous is immediately evident to them, and they can quickly proceed to apply the box-like setup or stereotyped procedure which they learned to thrust upon this "type." Apropos of this, we mention that our preliminary experiments indicated that (*a*) when the additional containers were very outstanding, either because of their size or position, the Einstellung effect was large and (*b*) when *S*'s received the basic setup followed by the four-jar variation, those who showed large

Einstellung effects in the former tended to do so even in the variation—they had learned the Einstellung method so well that no superfluous hypotheses stood in the way of their applying it.

To be effective, problems which aim at conveying the importance of discovering, selecting, evaluating, and discarding facts and hypotheses in solving problems, should be introduced in all school subjects and should not be treated as curiosities which must be heralded with a special introduction, but should be freely intermingled with other, more routine problems. If they involve insufficient or additional hypotheses, these should not be pattern as to number or kind. To be sure, as our experiments indicate, the inclusion of such problems may make learning slower and somewhat less efficient than drill procedures, but may also tend to produce less mechanical behavior and more productive thinking. Basically, it revolves on whether our schools wish to develop mechanical efficiency and a formula-applying attitude, conducive to associating a particular method with a particular situation, or whether they aim to develop individuals who have some capability in facing and coping with new and changing problem situations.

E. Concretizing the Tasks

We hoped that removing the problems from the abstract, symbolic level, and placing the emphasis on manipulation of jars rather than on written computations, might reduce the Einstellung effect. That there were *S*'s who regarded the numbers as essentially abstract symbols and who stressed computation involving the numbers, was seen when to the basic setup we added, as a last problem, one of the following: Given a 5-quart jar, a 25-quart jar, and a 10-quart jar, get 0 quarts; given a 3-quart jar, a 65-quart jar, and a 29-quart jar, get 3 quarts; given a 4-quart jar, a 67-quart jar, and a 17-quart jar, get 4 quarts. Fifty per cent of a class of college students blithely proceeded to obtain no water through the complicated mechanics of the Einstellung method: $25 - 5 - 2 \times 10 = 0$, quite indicative that this was to them an abstract, arithmetical operation. Thirty per cent of another college class (4, p. 73) devised an ingenious solution to the second-mentioned problem: $65 - 29 - 11 \times 3 = 3$; in "eleven times three" they repeatedly filled the 3-quart jar but failed to give the obvious solution to the problem—filling this jar once! Sixty-two per cent of a third college class (4, p. 73) failed to solve the last-mentioned problem, which required no complicated computations but merely the filling of the 4-quart jar. It seemed that these *S*'s did not view the figures as representations of the content of jars.

Finally, many participants in the basic experiment and its variations tended to transfer to the setup school-fostered attitudes toward arithmetic which appeared to favor Einstellung effect. We hoped that the effect of such attitudes would be diminished if in place of mere numbers actual jars were substituted, if the tasks were placed on a concrete level.

Outstanding differences between the basic setup and the one utilized in the present variation were that an actual supply of fluid was furnished, and that corresponding to the order in which the jars had been written in each problem, three containers were set on a table before *S*, with the corresponding volumes printed in large numerals on the surfaces facing *S*. The jars were made from used cardboard milk containers of the two-gill, pint, and quart size.[6] These were cut down so that the proportions were roughly similar in appearance to the proportions of the numerals labeled on them. To make the setup appear more realistic to *S*, the numerals were said to refer to the number of cubic centimeters which constituted the capacity of the jars, rather than to quarts as in the basic experiment.

The experiment was conducted in a classroom containing a sink; the sink faucet served as the source of water and the sink as the receptical for excess water. The setup was administered individually and was, except for the described differences, similar to individually conducted experiments with the corresponding arithmetical problems.

1. *Experiment I*

The experiment was conducted with 26 children of the sixth year of a New York City elementary school but only 22 solved the last two Einstellung problems, and only their results are considered. Pencil and paper were furnished at the outset "to be used if needed." All but two of the *S*'s first calculated the solution of each problem with pencil and paper before manipulating the jars. These two showed no Einstellung effects. The group as a whole gave Einstellung effects within the range of results made by similar groups which received the original setup: 68, 64, and 68 per cent Einstellung solutions of C_1C_2, #9, C_3C_4, respectively. Since so many *S*'s first figured out the solution on paper and then handled the jars in accordance with their calculations, our purpose in introducing the concrete jars was quite effectively defeated.

2. *Experiment II*

Thirty college students, all of whom solved the two tasks prior to the criticals in the Einstellung manner, gave as much Einstellung effect in C_1C_2 as did a similar group in the basic experiment, but showed better results in the remaining criticals. They gave 60, 33, and 33 percent Einstellung solutions of C_1C_2, #9, C_3C_4, respectively. Most of them regarded the entire affair as child's play until they were startled by the unsuccessfulness of their attempt to use the Einstellung method in Problem 9. Many first made some written calculations in each problem. Others quickly generalized the Einstellung method as a rule of solution and as soon as a set of three jars was presented, they rapidly filled the center one and poured from it, once into the container to the left and twice into the container to the right. It seemed that those who carefully examined each set of containers, who treated each problem as possessing individual requirements, and who used written calculations only as a check on their manipulations, showed little or no Einstellung effect.

[6]We also employed specially prepared glass containers blown up to the required dimensions but found that the jars made from milk containers, simpler to prepare, served the same purpose.

3. *Experiment III*

When pencil and paper were not provided and were expressly forbidden, 10 sixth-year elementary school pupils were found incapable of solving most of the Einstellung problems. Under these conditions, 20 college students, all of whom solved the two problems prior to the criticals in the Einstellung manner, made 55, 50, and 30 per cent Einstellung solutions of C_1,C_2, #9, C_3C_4, respectively. Some mentally calculated the arithmetic involved before manipulating the jars, thus putting the problem on a symbolic level. Others generalized the Einstellung method as a rule or formula, or claimed that they caught on to the "trick" which worked in these problems. Here too, there were differences in results between the foregoing and those who carefully surveyed each set of jars to see what method of solution it suggested.

4. *Discussion and Implications*

Concretizing the problems did not eliminate the Einstellung phenomenon. It appears that a tendency toward mechanization can occur both on the concrete and abstract levels. Thus the Einstellung tendency is not solely due to the fact that abstract symbols are involved in the basic setup. Of course it could have happened that what we witnessed in this experiment was the development, in the process of manipulating the actual jars, of a motor set, such as that found in weight lifting experiments. It is our impression, however, that this did not occur in most of the cases. The major factor in determining whether or not an Einstellung developed seemed to be the attitude with which *S* viewed the tasks.

One implication for education is that in teaching mathematics it is not sufficient to make the problems more concrete, more life like. The present trend toward concretizing mathematical problems by relating them more closely to everyday activities is in part motivated by the desire to make the subject matter more meaningful to the child; but this need not result in giving the child a better insight into mathematics—he may still repeat blindly certain rules and formulas. What are needed are teaching methods which will lead to understanding of the structural qualities of mathematical concepts and encourage productive thinking. (For a detailed discussion of such teaching methods the reader is referred to 2, 3, and 6.)

It seems to us that the use of actual jars furnishes a rich supply of information concerning the overt behavior of an *S* engaged in solving the volume-measuring problems and may broaden our understanding of the Einstellung phenomenon. Like the original setup, it seems to be an effective experiment for use in experimental psychology laboratories (cf. 5). It may be worthwhile to have students conduct both experiments and compare the observations and results. It may also be of value to conduct such variations of this setup as have been done with the numerical problems, e.g., to introduce speed conditions (4, pp. 53-56), to limit the quantity of available fluid, and to add a superfluous jar in each problem.

F. General Summary

We have previously reported (4, 5) upon extensive experimentation dealing with the tendency to impute an oft-repeated, rather complex mode of solution to subsequent problems which are capable of simpler solutions. This tendency toward mechanization has been taken as evidence of a mental set or Einstellung. The present report dealt with three new sets of experimental attempts to prevent or weaken the Einstellung tendency in solving a series of numerical problems dealing with volume-measurement.

In the first set of experimental variations an upper limit was placed on the amount of fluid available and *S*'s were cautioned to employ economical methods which did not waste the fluid; the non-habituated, simpler solutions were less wasteful of fluid. In the second series, an additional container was added to each problem so that the *S*'s could not rapidly impute the customary solution, but would first have to select the proper givens (or hypotheses) from those available. The third variation attempted to concretize the tasks by introducing actual jars and an actual supply of fluid in place of the abstract symbols usually employed.

On the whole, these experimental procedures were not very successful in decreasing the Einstellung effect, although in a few experiments statistically reliable differences in the expected direction were obtained. Often, the hoped for effects of the variations were vitiated by the carry-over to the experiment of school-learned attitudes toward arithmetic in particular and toward problem solving in general. For each method, educational implications were drawn.

References

1. Hilgard, E. R. Theories of Learning. New York: Appleton-Century-Crofts, 1948.
2. Duncker, K. On problem-solving. *Psychol. Monog.*, 1945, **58**, No. 270, 1-113. (Translated by Lynne S. Lees.)
3. Katona, G. Organizing and Memorizing. New York: Columbia Univ. Press, 1939.
4. Luchins, A. S. Mechanization in problem solving. *Psychol. Monog.*, 1942, **54**, No. 248, 1-95.
5. ———. Classroom experiments on mental set. *Amer. J. Psychol.*, 1946, **59**. 295-298.
6. Wertheimer, M. Productive Thinking. New York: Harper, 1945.

Department of Psychology
McGill University
Montreal, Canada