

THE INFLUENCE OF IMPROVEMENT IN ONE
MENTAL FUNCTION UPON THE
EFFICIENCY OF OTHER
FUNCTIONS. (I.)

BY DR. E. L. THORNDIKE,
Teachers College, New York,

AND DR. R. S. WOODWORTH,
New York University Medical School.

This is the first of a number of articles reporting an inductive study of the facts suggested by the title. It will comprise a general statement of the results and of the methods of obtaining them, and a detailed account of one type of experiment.

The word function is used without any rigor to refer to the mental basis of such things as spelling, multiplication, delicacy in discrimination of size, force of movement, marking *a*'s on a printed page, observing the word *boy* in a printed page, quickness, morality, verbal memory, chess playing, reasoning, etc. Function is used for all sorts of qualities in all sorts of performances from the narrowest to the widest, *e. g.*, from attention to the word 'fire' pronounced in a certain tone, to attention to all sorts of things. By the word improvement we shall mean those changes in the workings of functions which psychologists would commonly call by that name. Its use will be clear in each case and the psychological problem will never be different even if the changes studied be not such as everyone would call improvements. For all purposes 'change' may be used instead of 'improvement' in the title. By efficiency we shall mean the status of a function which we use when comparing individuals or the same individual at different times, the status on which we would grade people in that function. By other function we mean any function differing in any respect whatever from the first. We shall at times use the word function-group to mean

those cases where most psychologists would say that the same operation occurred with different data. The function *attention*, for instance, is really a vast group of functions.

Our chief method was to test the efficiency of some function or functions, then to give training in some other function or functions until a certain amount of improvement was reached, and then to test the first function or set of functions. Provided no other factors were allowed to affect the tests, the difference between the test before and the test after training measures the influence of the improvement in the trained functions on the functions tested.

It is possible to test the general question in a much neater and more convenient way by using, instead of measures of a function before and after training with another, measures of the correlation between the two functions. If improvement in one function increases the efficiency of another and there has been improvement in one, the other should be correlated with it; the individuals who have high rank in the one should have a higher rank in the other than the general average. Such a result might also be brought about by a correlation of the inborn capacities for those functions. Finding correlation between two functions thus need not mean that improvement in one has brought increased efficiency in the other. But the absence of correlation does mean the opposite. In an unpublished paper Mr. Clark Wissler, of Columbia University, demonstrates the absence of any considerable correlation between the functions measured by the tests given to students there. Miss Naomi Norsworthy, of Teachers College, has shown (the data were presented in part at the Baltimore meeting; the research is not yet in print) that there is no correlation between accuracy in noticing misspelled words and accuracy in multiplication, nor between the speeds; that there is little or no correlation between accuracy and speed in marking on a printed page misspelled words, words containing *r* and *e*, the word *boy*, and in marking semi-circles on a page of different geometrical figures.

Perhaps the most striking method of showing the influence or lack of influence of one function on another is that of testing the same function-group, using cases where there are very

slightly different data. If, for instance, we test a person's ability to estimate a series of magnitudes differing each from the next very slightly, and find that he estimates one very much more accurately than its neighbors on either side, we can be sure that what he has acquired from his previous experience or from the experience of the test is not improvement in the function-group of estimating magnitudes but a lot of particular improvements in estimating particular magnitudes, improvements which may be to a large extent independent of each other.

The experiments, finally, were all on the influence of the training on efficiency, on ability as measured by a single test, not on the ability *to improve*. It might be that improvement in one function might fail to give in another improved ability, but succeed in giving ability to improve faster than would have occurred had the training been lacking.

The evidence given by our experiments makes the following conclusions seem probable :

It is misleading to speak of sense discrimination, attention, memory, observation, accuracy, quickness, etc., as multitudinous separate individual functions are referred to by any one of these words. These functions may have little in common. There is no reason to suppose that any general change occurs corresponding to the words 'improvement of the attention,' or 'of the power of observation,' or 'of accuracy.'

It is even misleading to speak of these functions as exercised within narrow fields as units. For example, 'attention to words' or 'accurate discrimination of lengths' or 'observation of animals' or 'quickness of visual perception' are mythological, not real entities. The words do not mean any existing fact with anything like the necessary precision for either theoretical or practical purposes, for, to take a sample case, attention to the meaning of words does not imply equal attention to their spelling, nor attention to their spelling equal attention to their length, nor attention to certain letters in them equal attention to other letters.

The mind is, on the contrary, on its dynamic side a machine for making particular reactions to particular situations. It works in great detail, adapting itself to the special data of which

it has had experience. The word *attention*, for example, can properly mean only the sum total of a lot of particular tendencies to attend to particular sorts of data, and ability to attend can properly mean only the sum total of all the particular abilities and inabilities, each of which may have an efficiency largely irrespective of the efficiencies of the rest.

Improvement in any single mental function need not improve the ability in functions commonly called by the same name. It may injure it.

Improvement in any single mental function rarely brings about equal improvement in any other function, no matter how similar, for the working of every mental function-group is conditioned by the nature of the data in each particular case.

The very slight amount of variation in the nature of the data necessary to affect the efficiency of a function-group makes it fair to infer that no change in the data, however slight, is without effect on the function. The loss in the efficiency of a function trained with certain data, as we pass to data more and more unlike the first, makes it fair to infer that there is always a point where the loss is complete, a point beyond which the influence of the training has not extended. The rapidity of this loss, that is, its amount in the case of data very similar to the data on which the function was trained, makes it fair to infer that this point is nearer than has been supposed.

The general consideration of the cases of retention or of loss of practice effect seems to make it likely that spread of practice occurs only where identical elements are concerned in the influencing and influenced function.

The particular samples of the influence of training in one function on the efficiency of other functions chosen for investigation were as follows:

1. The influence of certain special training in the estimation of magnitudes on the ability to estimate magnitudes of the same general sort, *i. e.*, lengths or areas or weights, differing in amount, in accessory qualities (such as shape, color, form) or in both. The general method was here to test the subject's accuracy of estimating certain magnitudes, *e. g.*, lengths of lines. He would, that is, guess the length of each. Then he would

practice estimating lengths within certain limits until he attained a high degree of proficiency. Then he would once more estimate the lengths of the preliminary test series. Similarly with weights, areas, etc. This is apparently the sort of thing that happens in the case of a tea-taster, tobacco-buyer, wheat-taster or carpenter, who attains high proficiency in judging magnitudes or, as we ambiguously say, in delicacy of discriminating certain sense data. It is thus like common cases of sense training in actual life.

2. The influence of training in observing words containing certain combinations of letters (*e. g.*, *s* and *e*) or some other characteristic on the general ability to observe words. The general method here was to test the subject's speed and accuracy in picking out and marking certain letters, words containing certain letters, words of a certain length, geometric figures, misspelled words, etc. He then practiced picking out and marking words of some one special sort until he attained a high degree of proficiency. He was then re-tested. The training here corresponds to a fair degree with the training one has in learning to spell, to notice forms and endings in studying foreign languages, or in fact in learning to attend to any small details.

3. The influence of special training in memorizing on the general ability to memorize. Careful tests of one individual and a group test of students confirmed Professor James' result (see *Principles of Psychology*, Vol. I., pp. 666-668). These tests will not be described in detail.

These samples were chosen because of their character as representative mental functions, because of their adaptability to quantitative interpretations and partly because of their convenience. Such work can be done at odd times without any bulky or delicate apparatus. This rendered it possible to secure subjects. In all the experiments to be described we tested the influence of improvement in a function on *other functions closely allied to it*. We did not in sense-training measure the influence of training one sense on others, nor in the case of training of the attention the influence of training in noticing words on, say, the ability to do mental arithmetic or to listen to a metaphysical discourse. For common observation seemed to give a negative

answer to this question, and some considerable preliminary experimentation by one of us supported such a negative. Mr. Wissler's and Miss Norsworthy's studies are apparently conclusive, and we therefore restricted ourselves to the more profitable inquiry.

A SAMPLE EXPERIMENT.

There was a series of about 125 pieces of paper cut in various shapes. (Area test series.) Of these 13 were rectangles of almost the same shape and of sizes from 20 to 90 sq. cm. (series 1), 27 others were triangles, circles, irregular figures, etc., within the same limits of size (series 2). A subject was given the whole series of areas and asked to write down the area in sq. cm. of each one. In front of him was a card on which three squares, 1, 25 and 100 sq. cm. in area, respectively, were drawn. He was allowed to look at them as much as he pleased but not to superpose the pieces of paper on them. No other means of telling the areas were present. After being thus tested the subject was given a series of paper rectangles,¹ from 10 to 100 sq. cm. in area and of the same shape as those of series 1. These were shuffled and the subject guessed the area of one, then looked to see what it really was and recorded his error. This was continued and the pieces of paper were kept shuffled so that he could judge their area only from their intrinsic qualities. After a certain amount of improvement had been made he was retested with the 'area test series' in the same manner as before.

¹The judgments of area were made with the following apparatus: a series of parallelograms ranging from 10 to 140 and from 190 to 280 sq. cm., varying each from the next by 1 sq. cm. Their proportions were almost the same (no one of them could possibly be distinguished by its shape). For example, the dimensions of those from 137 to 145 sq. cm. were

| | |
|-------|----------|
| 15 | × 9.133½ |
| 15 | × 9.2 |
| 15.1 | × 9.2 + |
| 15.1 | × 9.275 |
| 15.15 | × 9.31 |
| 15.15 | × 9.375 |
| 15.2 | × 9.41 |
| 15.2 | × 9.475 |
| 15.2 | × 9.54 |

The function trained was that of estimating areas from 10 to 100 sq. cm. with the aid of the correction of wrong tendencies supplied by ascertaining the real area after each judgment. We will call this 'function *a*.' A certain improvement was noted. What changes in the efficiency of closely allied functions are brought about by this improvement? Does the improvement in this function cause equal improvement (1) in the function of estimating areas of similar size but different shape without the correction factor? or (2) in the function of estimating identical areas without the correction factor? (3) In any case how much improvement was there? (4) Is there as much improvement in the function of estimating dissimilar shapes as similar? The last is the most important question.

We get the answer to 1 and part of 3 by comparing in various ways the average errors of the test areas of dissimilar shape in the before and after tests. These are given in Table I. The average errors for the last trial of the areas in the training series similar in size to the test series are given in the same table.

TABLE I.

| Subject | Test series 2 | | Training series Av error at end of training |
|---------|------------------------------|----------------------------|---|
| | Av. error before training | Av error after training | |
| T | 15.8 | 11.1 | 2.3 |
| Be. | 28.0 | 5.2 | 3.1 |
| Br. | 22.5 | 18.7 | 3.3 |
| J. W. | 12.7 | 21.0 | 1.5 approx. |
| W (2) | 17.0 | 20.0 | 4.0 " |
| E. B. | 10.5 | 7.9 | 0.4 |

The function of estimating series 2 (same sizes, different shapes) failed evidently to reach an efficiency equal to that of the function trained. Did it improve *proportionately* as much?

This is a hard question to answer exactly, since the efficiency of 'function *a*' increases with great rapidity during the first score or so of trials, so that the average error of even the first twenty estimates made is below that of the first ten, and that again is below that of the first five. Its efficiency at the start depends thus on what you take to be the start. The fact is that the first estimate of the training series is not an exercise of 'function *a*' at all and that the *correction* influence increases

up to a certain point which we cannot exactly locate. The fairest method would seem to be to measure the improvement in 'function *a*' from this point and compare with that improvement in the other function or functions in question. This point is probably earlier in the series than would be supposed. If found, it would probably make the improvement in 'function *a*' greater than that given in our percentages.

The proportion of average error in the after test to that in the before test is greater in the case of the test series than in the case of the first and last estimations of the areas of the same size in the training series, save in the case of *Be*. The proportions are given in the following table :

TABLE II.

| Proportion of error after to error before training | | |
|--|---------------|-----------------|
| Subject | Test series 1 | Training series |
| T. | .70 | .575 |
| Be. | .19 | .56 |
| Br. | .83 | .53 |
| J. W. | 1.75 | .77 approx. |
| W (2) | 1 18 | .83 approx |
| E. B. | .75 | 13 |

Question 2 is answered by a comparison of the average errors, before and after the training, of Series I. (identical areas) given without the correction factor. The efficiency reached in estimating without the correction factor (see column 2 of Table III.) is evidently below that reached in 'function *a*.' The results there in the case of the same areas are given in column 3.

TABLE III.

| Subject | Av. error before training of series 1 | Av error after training of series 1 | Av error after training of same sizes in training series. | Av. error after training of series 2 | Proportion of error after to error before training | | |
|---------|---------------------------------------|-------------------------------------|---|--------------------------------------|--|-----------|---|
| | | | | | Series 2 | Series 1. | Areas of training series identical with series 1. |
| T. | 9.0 | 6.0 | 2.1 | 11.1 | .70 | 67 | .31 |
| Be. | 21.9 | 6.4 | 1.8 | 5.2 | .19 | .29 | .45 |
| Br. | 24.2 | 14.7 | 3.7 | 18.7 | .83 | .61 | .37 |
| J. W. | 7.7 | 8.6 | 1.5 app. | 21.0 | 1.75 | 1.11 | .77 app. |
| W. (2) | 11.6 | 3.3 app. | 4.0 app. | 20.0 | 1.18 | .28 app. | .83 app. |
| E. B. | 9.8 | 4 1 | 0.4 | 7.9 | .75 | .42 | .08 |

The function of estimating an area while in the frame of mind due to being engaged in estimating a limited series of areas and seeing the extent of one's error each time, is evidently independent to a large extent of the function of judging them after the fashion of the tests.

If we ask whether the function of judging without correction improved proportionately as much as 'function *a*,' we have our previous difficulty about finding a starting point for *a*. Comparing as before the first 100 estimates with the last 100 we get the proportions in the case of the areas identical with those in the test. These are given in column 7. The proportions in the case of the test areas (series 1; same shape) are given in column 6. A comparison of columns 6 and 7 thus gives more or less of an answer to the question, and column 6 gives the answer to the further one: "How much improvement was there?"

We can answer question 4 definitely. Column 5 repeats the statement of the improvement in the case of the test areas of different shape, and by comparing column 6 with it we see that in every case save that of Be. there was more improvement when the areas were similar in shape to those of the training series. This was of course the most important fact to be gotten at.

To sum up the results of this experiment, it has been shown that the improvement in the estimation of rectangles of a certain shape is not equalled in the case of similar estimations of areas of different shapes. The function of estimating areas is really a function-group, varying according to the data (shape, size, etc.). It has also been shown that even after mental standards of certain limited areas have been acquired, the function of estimating with these standards constantly kept alive by noticing the real area after each judgment is a function largely independent of the function of estimating them with the standards fully acquired by one to two thousand trials, *but not constantly renewed by so noticing the real areas*. Just what happened in the training was the partial formation of a number of associations. These associations were between sense impressions of particular sorts in a particular environment coming to a person in a particular mental attitude or frame of mind, and a number of ideas or impulses.

What was there in this to influence other functions, other processes than these particular ones? There was first of all the acquisition of certain improvements in mental standards of areas. These are of some influence in judgments of different shapes. We think, "This triangle or circle or trapezoid is about as big as such and such a rectangle, and such a rectangle would be 49 sq. cm." The influence is here by means of an idea that may form an identical element in both functions. Again, we may form a particular habit of making a discount for a tendency to a constant error discovered in the training series. We may say, "I tend to judge with a minus error," and the habit of thinking of this may be beneficial in all cases. The habit of bearing this judgment in mind or of unconsciously making an addition to our first impulse is thus an identical element of both functions. This was the case with *Be*. That there was no influence due to a mysterious transfer of practice, to an unanalyzable property of mental functions, is evidenced by the total lack of improvement in the functions tested in the case of some individuals.

On pushing our conception of the separateness of different functions to its extreme, we were led to ask if the function of estimating one magnitude might not be independent even of the functions of estimating magnitudes differing only slightly from the first. It might be that even the judgment of areas of 40-50 sq. cm. was not a single function, but a group of similar functions, and that ability might be gained in estimating one of these areas without spreading to the others. The only limits that must necessarily be set to this subdivision would be those of the mere sensing of small differences.

If, on the contrary, judgments of nearly equal magnitudes are acts of a single function, ability gained in one should appear in the others also. The results of training should diffuse readily throughout the space covered by the function in question, and the accuracy found in judgments of different magnitudes within this space should be nearly constant. The differences found should simply be such as would be expected from chance.

The question can be put to test by comparing the actual difference between the average errors made, in judging each of

neighboring magnitudes, with the probable difference as computed from the probability curve. If the actual difference greatly exceeds the probable difference, it is probably significant of some real difference in the subject's ability to judge the two magnitudes. He has somehow mastered one better than the other. No matter how this has come about. If it is a fact, then clearly ability in the one has not been transferred to the other.

Our experiments afford us a large mass of material for testing this question. In the 'training series,' we have a considerable number (10 to 40) of judgments of each of a lot of magnitudes differing from each other by slight amounts. We have computed the accuracy of the judgment of each magnitude (as measured by the error of mean square), and then compared the accuracy for each with that for the adjacent magnitudes. We find many instances in which the difference between the errors for adjacent magnitudes is largely in excess of the probable difference. And the number of such instances greatly exceeds what can be expected from chance.¹

These great differences between the errors of adjacent magnitudes are strikingly seen in the curves on page 259. The ordinates of these curves represent the mean square error of judgments of areas of 10 to 100 square centimeters, and for 3 individuals. The dots above and below each point of the curve give the 'limits of error' of that value, as determined by the formula, $\frac{\mu}{\sqrt{2n}}$, in which μ is the error of mean square, and n the number of cases. These limits are such that the odds are about 2 to 1, more exactly 683 to 317, that the true value lies inside them. The dots thus furnish a measure of the reliability of the curve at every point.

¹ The smaller error at certain magnitudes is not the result of a preference of the subject to guess that number. Of course, if the subject were prone to guess '64 square centimeters' oftener than 63 or 65, he would be more apt to guess 64 right, and the error for 64 would be diminished. We therefore made a few tables of the frequency with which each number was guessed. But we found that the magnitudes that were best judged were not more often guessed than their neighbors.

These curves are all irregular, with sudden risings and fallings that greatly obscure their general course. Psychologists are familiar of old with irregularities of this kind, and are wont to regard them as effects of chance, and so to smooth out the curve. But as we find more irregularity than can reasonably be attributed to chance, we conclude that our curves at least should not be smoothed out, and that the sudden jumps, or some of them, signify real differences in the person's ability.¹

If, for example, we examine Fig. 1, we notice a number of sudden jumps, or points at which the errors in judging adjacent magnitudes differed considerably from each other. The most significant of these jumps are at 10-11, 36-37, 41-42, 65-66, 66-67, 83-84, and 98-99 sq. cm. The question is whether such a jump as that at 41-42 indicates greater ability to judge 42 sq. cm., or whether the observed difference is simply due to chance and the relatively few cases (here 10 for each area). A vague appeal to chance should not be allowed, in view of the possibility of calculating the odds in favor of each side of the question. This can be done by a fairly simple method. We can consider two adjacent areas as practically equal, so far as concerns Weber's law or any similar law. The average errors found for the two would thus be practically two determinations of the same quantity, and should differ only as two determinations of the same quantity may probably differ.

We wish then to compare the actual difference between the errors for 41 and 42 sq. cm. with the probable difference. The error—we use throughout the 'error of mean square,' and the measure of reliability based on it—this error is here 6.2 and 3.1 sq. cm. respectively. The actual difference is 3.1 sq. cm. To find the probable difference, we first find the 'limits of error' or reliability of each determination, as described above, and then find the square root of the sums of the squares of these

¹The fact that judgments of nearly equal magnitudes may show very unequal errors throws doubt on all curves drawn from the judgment of only a few 'normals.' If slightly different normals had been chosen, the errors might have been considerably different, and the course of the curve changed. If, for example, three normals be chosen from the 91 in our curves, and those three used as the basis of a curve, the curve will vary widely with the choice of the three normals.

FIG. 1.

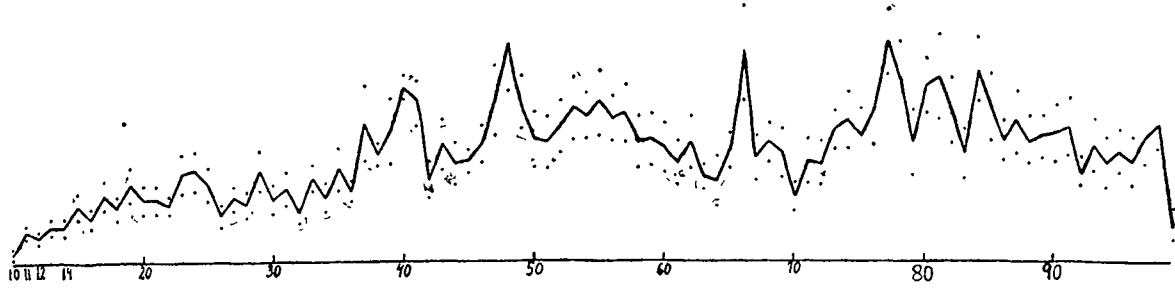


FIG. 2.

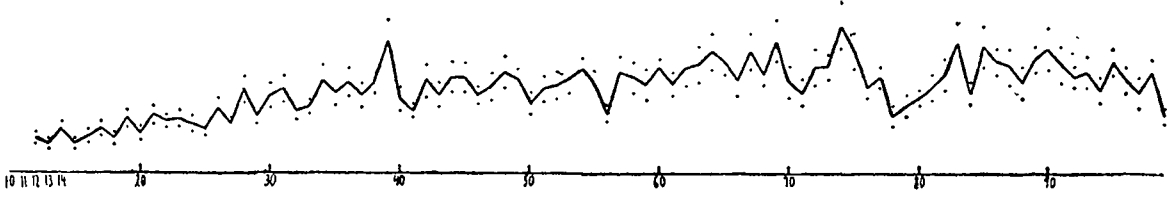
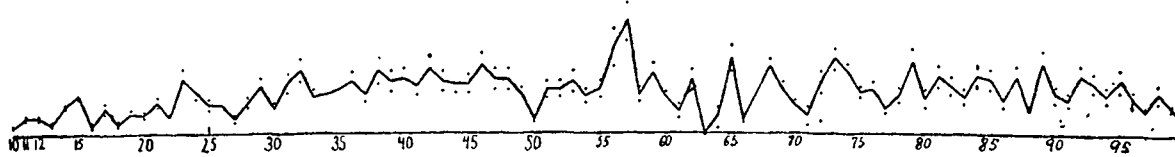


FIG. 3.



'limits of error.' The 'limits' are here 1.0 and 0.7, and the probable difference 1.2 sq. cm. The actual difference is 2.6 times the probable. In this whole series we find 6 other instances in which the actual difference is over 2 times the probable. From the probability integral we find that, in the long run, 46 actual differences to the thousand would exceed twice the probable. The question is, therefore, what is the probability of finding as many as 7 such differences in a series of 90? This is a form of the familiar problem in probabilities: to find the chances that an event whose probability is p shall occur at least r times out of a possible n . The solution depends on an application of the binomial theorem, and may be evaluated by means of logarithms. In the present case, the value found is .1209 or about $\frac{1}{8}$.

Instead of vaguely saying that the large jumps seen in the curves may be due to chance, we are now able to state that the odds are 7 to 1 against this view, and 7 to 1 in favor of the view that the large jumps, or some of them, are significant of inequality in the person's power to estimate nearly equal areas. These odds are of course not very heavy from the standpoint of scientific criticism. But they are fortified by finding, as we do, the same general balance of probability in all of the series examined. In one other series, the number of large differences is small, and the probability is as large as .2938 that they are due to mere chance. In three other series, this probability is very small, measuring .0025, .0025, .0028, or about $\frac{1}{400}$. Finally, in the series corresponding to Fig. 3, there are a large number of actual differences which far exceed the probable. (The errors are small, and consequently the probable differences are small.) There are 31 that exceed twice the probable difference, and of these 9 exceed 3.5 times the probable difference. The probability of finding even these 9 is so small that six-place logarithms cannot determine it exactly, but it is less than .000001.

In four cases, then, out of six examined, it is altogether inadmissible to attribute the differences to chance, while in the other two the odds are against doing so. The probability that the differences in *all* the series are due to chance is of course

multiply small. The differences are therefore not chance, but significant; the ability to judge one magnitude is sometimes demonstrably better than the ability to judge the next magnitude; one function is better developed than its neighbor. The functions of judging nearly equal magnitudes are, sometimes at least, largely separate and independent. A high degree of ability in one sometimes coexists with a low degree of ability in the others.