ACCURACY IN SCHOOL CHILDREN. DOES IMPROVEMENT IN NUMERICAL ACCURACY "TRANSFER"?

W. H. WINCH,

London, England.

SUMMARY.

Pupils in four different schools were studied with reference to the question, "Does improvement in accuracy of numerical computation 'transfer' to arithmetical reasoning?" The pupils of each school were divided into two equal groups on the basis of preliminary tests in working arithmetical problems. One of these groups was practised in working rule sums, while the other was occupied with history, drawing, etc. After a certain number of days of practise further sets of problems were used to test both groups in arithmetical reasoning. In the first school tested there was no evidence of 'transfer,' but in the second, third and fourth schools the practised section showed certain improvements in arithmetical reasoning.

Many more experiments of this sort are necessary before definite conclusions can be drawn, but from the present investigation it would seem that some sort of connection may exist between improvement in numerical computation and mathematical reasoning. The results are too irregular to warrant the conclusion of any definite 'transfer,' but the improvement in reasoning may be due either to release of mental energy resulting from improved facility in computation, or to the association established between the two kinds of function usually at work together, whereby continuance of the one involves greater readiness for the other. It is hoped that others will be stimulated to repeat the tests and to carry them farther.

I. INTRODUCTION.

It is a reproach frequently made by educationists against psychologists that the latter do not apply themselves to the

*The experiments here described in detail were given in outline in a paper on the "Faculty Doctrine," read at a joint meeting of the Aristotelian and British Psychological Societies and of the *Mind* Association. in London, July, 1910.

solution of problems which have direct and immediate bearings on educational theory and practice; and it is further complained that, however true many psychological results may be, they are nevertheless, irrelevant, as far as education is concerned. Perhaps some justification for the complaint may be conceded; but the educationist, on his side, may in fairness be required to remember that educational psychology is not the whole of psychology, and further, that no science can, or ought to, be expected to issue in practical directions at every step of its growth and progress.

II. THE PROBLEM STATED.

Among the practical questions which are hotly debated by educationists today, there are to be found problems whose solution will not only be of importance to the educator, but may throw some light also upon psychological and epistemological questions—such an one is that of “accuracy.”

i. *Its pedagogical aspect.*

It is impossible for a visitor to ask questions in schools as to the *rationale* of methods and processes adopted by teachers without continually receiving answers which imply the “faculty doctrine” in its extremest form. If one asks, for example, why a high standard of numerical accuracy is demanded in arithmetical exercises, it is relatively rare that one receives an answer based upon the practical inconveniences which arise from inaccurate results.

It is much more common to be told that children ought to be accurate in arithmetic because the *habit* of accuracy is so important. And when one asks why a much higher standard of proficiency in numerical computation is demanded in school than, in most cases, life is likely to require, an answer is given somewhat in the following form, “Well, you see, it makes the child accurate all round.” If, at this point, one asks whether improvement in numerical accuracy will improve accuracy in drawing or accuracy in the sight reading of music, there is apt to be some hesitation and search-

ing of heart. But if the question be put, "Will it make the boy more accurate in other arithmetical operations?" the answer is unhesitatingly in the affirmative. This then is our problem today. Is there such a thing as a habit of accuracy? Is accuracy a "faculty"? Will accuracy, developed by training the child in one school subject, be transferred to another in which he has not been trained? Will it be transferred even to other aspects of the same school subject, e. g., arithmetic? May it not be found that even this well-known "subject" is not so much one subject as a congeries of independent operations?

That these questions are not superfluous at the present time may be seen by the alarming variety in the views held by competent educationists. Mr. J. L. Paton, in an article in the *Educational Times* for July, 1907, writes accuracy as one of the school virtues which are in reality great life virtues exemplified and enforced in small matters—the implied argument being that accuracy in school work will tend to accuracy in the work of life, whatever that may be, and *à fortiori*, that accuracy in one branch of school work will be reflected in the accuracy of another.

The opposed view is very strongly put in Chapter VIII. of Dr. F. H. Hayward's *Education and the Heredity Spectre*; he says that the "dogma of formal training—faith in which constitutes the greatest educational superstition of today—needs to be statistically *refuted*" (italics mine). It would be better, perhaps, to ask that the doctrine might be statistically tested, for it sometimes happens that figures collected to overthrow a doctrine are found to establish it.

If these discussions and differences were to be found only in academic realms and in polemic works of small circulation they would be of relatively small importance; but that is not the case. The inspector and the school-master take a hand, especially in arithmetical questions, either asserting that improvement in mechanical numerical accuracy will affect a boy's arithmetical reasoning powers adversely, or

that accuracy in one will lead to accuracy in the other.

It not unfrequently happens that both parties to the dispute—when experimental inquiry is made—are found to be right; it sometimes happens that they are both found to be wrong. I propose, however, to leave all inferences until after the method and results of the experimental work have been presented.

ii. *Its psychological aspect.*

The grouping of mental operations into classes has always been found convenient, and is likely, notwithstanding animadversions about 'faculty psychology,' to go on. Is 'accuracy' a generic name which may fairly be applied in this sense? Is accuracy a faculty? Is there a habit of accuracy involving more than the operations whose accuracy has been increased by training?

iii. *Its epistemological aspect.*

What relationship must be shown to exist between two powers or functions of the same mind before we are entitled to call them part of the same 'faculty' or group of similar functions? Suppose we find high positive correlation between two mental operations, what sort of relationship may be inferred between them? Have we a guarantee of any relationship beyond what a metaphysician might call a mere "togetherness"?

The educationist, let us remember, when he hears psychologists and others talk of correlations, thinks that the establishment of high positive correlations means much more than this. It is useful, doubtless, for a teacher, especially at examination times, to be able to tell from a test in one or two subjects or parts of subjects what the class will be able to do in others. Mere "togetherness," if quantified, will enable this to be done. And even that is not without value.

But what the educationist is apt to assume high positive correlation to mean—and not only the educationist—is that a relationship exists between two functions of such a nature

that improvement in the one implies improvement in the other (subject to certain conditions as to fatigue and saturation point), and that decay of the one implies decay of the other.

For if there is not causal dependency, it is frequently thought that there is some central or common factor to whose existence the positive correlation is due, the training of which would therefore be common to the two operations.

III. GENERAL PLAN OF THE EXPERIMENTS.

The general method adopted was the same throughout, though there were minor modifications as to the number and difficulty of the exercises in different schools. In each case a whole class, working according to the same syllabus of work and under one teacher, was divided into two equal groups. The division was effected on the results of several tests in problematical arithmetic. In order that the natural ability of the children rather than their memory of recent teaching should be tested in these exercises, it was arranged that no problems should be given in a form with which the pupils were well acquainted. The tests were marked solely with reference to the accuracy of the arithmetical reasoning and entirely without reference to the accuracy of the numerical computation. No attention was paid to the *right answer*; marks were given with reference to the process only.

When the two equal groups had been obtained, one of them was practised in a series of exercises in 'rule' sums which every child knew how to do; the other group being meanwhile engaged in some other branch of school work. In every other respect the curriculum for both groups was precisely the same during the period of the experiment. I need, perhaps, hardly say that no other arithmetical work was done during the time the experiment lasted.

Finally, the two groups, namely, the one practised in accurate numerical computation and the one not so practised, were placed together again, and final tests were given in

arithmetical reasoning. There had been, in all cases, a considerable improvement in the accuracy of numerical computation during the series of practice exercises. How far was this improvement of accuracy in arithmetical computation transferred to accuracy in arithmetical reasoning? Did the practised or non-practised groups do better work when the groups worked the same test exercises at the end?

The above issue is, pedagogically, the most important of all those on which the following experiments throw light; it would be referred to in educational works as the question of the transfer of improvement to accuracy.

These experiments, however, help us towards the answer to another question which is primarily epistemological though of great importance to psychology. Let us find the correlation between the two forms of accuracy — accuracy of arithmetical reasoning and accuracy of numerical computation. If it is found to be highly positive, we shall suppose either that one function causes the other, or that there is a common factor in the two functions, or that both functions are the result of some prior quality or power which develops approximately equally in both these directions, or, finally, that the high correlation is a mere togetherness without relations or interdependence.

The two former hypotheses are capable of verification by means of our experiments. If one causes the other, then improvement in one should lead to improvement in the other. If there is a common factor, improvement in one should (so far as the common factor goes) lead to improvement in the other.

To find the correlation between accuracy in numerical computation and accuracy in arithmetical reasoning, it would seem, at first, a satisfactory procedure to correlate the marks for the preliminary problematic lists and the practice exercises in rule sums which were worked subsequently by one of the two equal groups. But it would be unsatisfactory to correlate two series obtained under dif-

ferent conditions as to time and curriculum. If one desires to find a correlation which shall be free from these disturbing influences, it is better to take the exercises on which it is to be based side by side and under the same school conditions. This, therefore, was done. Tests in rule sums, which every child knew *how* to do, were given *pari passu* with the tests in arithmetical reasoning. So that the conditions as to fatigue (if any) might not be more favorable to one sort of test than to the other, in one school the problem and rule sums were given alternately first and second. This variation produced no change in the results; it appeared to be an unnecessary precaution.

Having found the correlation between the two functions which, I may say in advance, is highly positive, we may refer to the tables which show the transfer (if any) of improvement in accuracy from one function to the other. If the result is negative, small, or doubtful, we are, on empirical grounds, faced by very serious limitations to the inferences we may draw from the fact of high positive correlation between two functions.

A second point, rather, perhaps, of pedagogical than of psychological interest, is the following: Besides the most important issue, namely, the decision as to the question of the transfer of improvement in the accuracy of numerical computation to accuracy in problematic reasoning, there is the subsidiary one: How far is the improvement in numerical accuracy gained during the practise exercises transferred to the numerical computation within the final sets of problems worked? For it is frequently thought, even by competent opinion, that practise work in accuracy of this kind is practically useless when the pupil sets to work on problematic sums.

IV. FIRST SERIES OF EXPERIMENTS. SCHOOL "S."

A first series of experiments was carried out in a municipal girls' school situated in a poor neighborhood, with Standards VI. and VII. of an average age at the beginning of the

series of 13 years 2 months. The teacher of the class was disposed to throw emphasis on the problematic rather than the computational side of arithmetic. The problems she was accustomed to give were difficult for this type of child and the tests set are hard. I mention this because the schools selected by me to do this work had different attitudes towards arithmetical teaching, for I was anxious to avoid finding correlations and other results which might have been dependent rather on the trend of the teaching than on the natural abilities of the children.

I avoid dealing with the question as to how far the trend of the teaching *can* obscure or overcome natural correlations. I am of opinion — an opinion not widely shared — that its power is rather limited, but I thought it best to vary my examples in this respect so as to strengthen my conclusion if the results appeared unaffected by the educational *milieu*.

This work was done in the spring of 1909. A brief chronology of the experiment with some specimens of the tests and exercises follows:—

Tuesday, March 30th.

10.15 A. M. to 10.35 A. M. All the classes worked arithmetical problems.

10.40 A. M. to 11.00 A. M. All the class worked rule sums.

Thursday, April 1st.

10.15 A. M. to 10.35 A. M. All the class worked arithmetical problems.

10.40 A. M. to 11.00 A. M. All the class worked rule sums.

Friday, April 2nd.

10.15 A. M. to 10.35 A. M. All the class worked arithmetical problems.

10.40 A. M. to 11.00 A. M. All the class worked rule sums.

The class was now divided into two equal groups on the basis of the marks for reasoning obtained in working the problem papers.

Tuesday, April 6th.

10.15 A. M. to 10.55 A. M. One of the two groups — hereafter called the practised group — worked rule sums. The other group studied history.

At this stage of the experiment the Easter holiday intervened, after which the exercises were resumed.

Tuesday, April 20th.

10.15 A. M. to 10.55 A. M. The practised group worked rule sums:
the non-practised group did drawing.

Thursday, April 22nd.

10.15 A. M. to 10.55 A. M. The practised group worked rule sums:
the non-practised group did drawing.

Friday, April 23rd.

10.15 A. M. to 10.55 A. M. The practised group worked rule sums:
the non-practised group did drawing.

The two groups were now placed together again, and on

Tuesday, April 27th.

10.15 A. M. to 10.55 A. M. All the class worked a test in problematic arithmetic; a longer paper of questions was set on this occasion than during the preliminary problematic tests.

i. *Specimens of Tests and Exercises with the Method of Marking.*

(a.) *A Set of Arithmetical Problems.*

1. A certain grocer gains $\frac{3}{4}$ d. on every 4 ounces of coffee which he sells; what would he gain by the sale of $1\frac{1}{2}$ cwts.?
2. A woman earns three times as much as a boy and a man twice as much as a woman. If the daily wages of a man, a woman and a boy amount to 19s. 2d. what are the daily earnings of each?
3. A tank is full of water. After $\frac{1}{3}$ of its contents has been drawn off, $\frac{1}{5}$ of the remainder is also drawn off. If 18 gallons are then left in the tank, what does it hold when full?

As already stated these tests were marked so as to give credit for arithmetical *reasoning* only, irrespective of the correctness of the answer numerically.

But a mere glance will show that not all the above sums are equal in difficulty. How shall we mark so that our units shall be approximately equal and that the obviously more difficult sum shall receive an adequately higher mark? I suggest the following method with some confidence (though I do not consider it perfect), as I have worked with it for some years with satisfactory results.

Let us consider each necessary rational 'step' as worthy of *one* mark. In the first problem, for instance, it is essential to the solution (at the mental stage reached by these children), (1) that the ounces and the hundredweights should be reduced to the same denomination, either the ozs. to cwts. or the cwts. to ozs.; (2) that the number of times which 4 ounces is contained in the total weight should be found; (3) that the gain on each 4 ounces should be multiplied by the quotient thus obtained. This gives us a total of 3 rational steps or marks for the first problem.

The second problem is much more difficult. To solve this the children require to regard (1) the boy as possessing one share, (2) the woman as possessing three times as much as the boy, namely, three shares, (3) the man as possessing six times as much as the boy, namely, six shares. Then (4) the shares must be added, and (5) the 19s. 2d. must be divided by the total number of shares. The quotient is the boy's wage. Thus we have a total of 5 marks for reasoning for the second problem. The multiplication of the boy's wages is regarded as rationally identical with the earlier steps, so no additional marks are awarded.

The third problem is resolved by the following rational steps: (1) The $\frac{1}{3}$ drawn off must be subtracted from unity. (2) $\frac{2}{3}$ of the remainder must be found. (3) The total fraction drawn off must be found. (4) The fraction of the whole now remaining must be calculated. (5) The fraction must be equated with the gallons left in the tank. (6) From the value of the given fraction in gallons the value of unity in gallons must be calculated. This gives a total of six marks for the rational steps of the third problem.

I need hardly say that any rational step necessary towards, or on the way towards, a correct solution would have received a mark; but, in practice, the marks actually obtained by the pupils of this class were obtained for some or all of the above 'steps'. It is not maintained that this system of marking is theoretically perfect, but it gives steady

and reliable results, and is a system which teachers readily appreciate and understand. This latter condition seems to me one of the utmost importance in all research work on a large scale in educational psychology.

(b) *A Test in Rule Sums worked immediately after a Test in Arithmetical Problems.*

- (1). Divide 3 cwt. 2 qrs. 21 lbs. 7 ozs. by 35.
- (2). Divide $\frac{1}{2} \times \frac{1}{3} - \frac{1}{4} \div 3$ 1-23.
- (3). Express 1 cwt. 2 qrs. as the fraction of 2 cwt. 3 qrs. 4 lbs.

(c) *A Practical Exercise in Rule Sums given to the practised group only.*

- (1). £1398 7s. 0 $\frac{3}{4}$ d. $\div$ 39.
- (2). Take £77 2s. 6 $\frac{3}{4}$ d. from £1000 and multiply the answer by 25.
- (3). What is the value of 7-9 of £6 18s. 9d.
- (4). 3 $\frac{1}{4}$ of 5-7 of 9 3-15 of 3 $\frac{1}{2}$.
- (5). $\frac{5}{8}$ of a ton $+$ 3-7 of a quarter.
- (6). $(.007 + 5.01 - 1.05) \times .068$.

The method of marking adopted for 'accuracy' in computation, though extremely laborious, was, in principle, extremely simple. It was intended that no sum should be set in these exercises which the children did not know *how* to do; but of course, though the processes were well known, there was much variation in accuracy of computation. It is obvious to any one who has kept records of arithmetical work from week to week that the usual method of marking sums 'right' or 'wrong' and allowing for each sum so many marks, all or most of which are lost if there is any numerical inaccuracy, will not give us a series of results which will adequately show improvement in accuracy. A child with one or two figures wrong in several sums may really have worked more accurately than another child who has all the sums right but one, provided that that one be wrong in many places.

Further, a child may make much improvement, even though its total number of sums right does not increase. Having these considerations in mind it was decided to give

one mark for each correct unit of addition, subtraction, multiplication and division. Thus, if £1398 7s. 0¾d. be divided by 39, the first unit is a unit in division. A three in the quotient receives one mark. The multiplication of 39 by 3 receives three marks; thus, 3 times 9 are 27, the seven receives a mark; 3 times 3 are 9, the 9 receives a mark, and if the + 2 from the 27 is added correctly (teachers know how often 'carried numbers' are added incorrectly) another mark is given. The subtraction of 117 from 139 is assessed in a similar way; each unit process receives one mark. Some—a very few—able children telescope processes which average children work out at greater length, but their units are always calculated on the most generous basis. A child on this method of marking does not lose all the rest of the marks for that sum because one number early in the sum is found wrong, even though subsequent operations on that number lead to further errors in the sum as usually understood. He may not really have made any *more* errors; it is necessary to go through the sum, unit by unit, to see if he has or has not. Doubtless some of my readers have said to themselves, "Why not use 'reckoning books', where the unit of calculation would be so much more easily estimated?" I did not do this, since it would have introduced an element of disturbance in the normal work of the children; and further, because teachers are not usually convinced by experiments which seem to them something outside and apart from ordinary school-work.

ii. *Results.*

The marks for accuracy in arithmetical reasoning in the three preliminary tests were added together, and the marks for accuracy in numerical computation in the tests worked side by side (immediately after) with the problems in reasoning were also added together. The results follow:

Table I., showing the general correlation between
 (a) *accuracy in reasoning,*
 (b) *accuracy in numerical computation.*

Marks for Reasoning.	No. of Children.	Average Mark for Reasoning Per Child	Av Mark for Numerical Computation Per Child.
Over 30	2	32.5	306.0
25 to 30	9	28.0	282.6
20 to 25	4	21.7	284.2
15 to 20	9	18.2	267.2
10 to 15	8	11.6	224.0

It is fairly obvious from even a brief inspection of the above table that children who are good at arithmetical reasoning are usually good at numerical computation also, i. e., the correlation is positive. It is wise, however, to find a more definite value than can be obtained from the table. The children were placed on a list thus:

Name (Initials Only)	Marks for Reasoning	Marks for Computation (To the nearest ten.)
1. A. P.	33	31
2. E. C.	32	30
3. A. B.	30	29
.....	..	..
32. R. D.	10	20

Then the Pearson formula $r = \frac{\sum xy}{n o_1 o_2}$ was applied to the individual cases and 'r' was found to be + .68.

May we from this result presume that the two powers, accuracy in arithmetical reasoning and accuracy in numerical computation, are so related that improvement in one of them may be found to result in improvement in the other, that is, that accuracy transfers? Let us proceed to test the matter empirically.

After the material for correlation had been obtained the class was divided into two equal groups on the basis of the marks obtained in the three preliminary tests of arithmetical reasoning.

The result follows:

Table II., showing the division into two equal groups.

GROUP A.		GROUP B.	
Name (Initials Only.)	Marks for Reasoning.	Name (Initials Only)	Marks for Reasoning.
A. P.	33	E. C.	32
L. F.	28	A. B.	30
D. W.	28	E. R.	30
E. R.	28	J. B.	28
M. D.	27	L. M.	27
M. C.	26	A. W.	24
E. A.	21	D. Hd.	21
A. R.	20	D. Hn.	21
M. J.	19	M. R.	19
R. S.	19	C. B.	19
W. O.	18	E. W.	18
E. W.	16	E. G.	16
L. J.	14	F. D.	13
M. O'H.	12	J. B.	12
D. G.	11	A. H.	11
L. P.	10	R. D.	10

The group selected to do the practice exercises in 'rule sums' was determined by 'chance', which fell upon Group B, hereafter called the 'practised' group. A record of units of accuracy in computation was kept during the course of the practice exercises. The improvement shown by the practised group in numerical accuracy may be seen from the following table:

Table III., showing the improvement in numerical accuracy of the practised group.

Marks for Preliminary Tests in Reasoning.	No of Children.	Marks for Numerical Accuracy in the Practice Exercises			
		Av 1st Ex.	Av. 2nd Ex	Av. 3rd Ex.	Av. 4th Ex.
30 and over.	1	196	207	207	251
25 to 30	4	178	195	214	233
20 to 25	3	218	219	232	218
15 to 20	4	171	181	193	177
10 to 15	4	145	169	145	128

In the fourth practice exercise one sum was set, which D. Hn. in the third section, E. W. in the fourth section, and

J. B. in the fifth or lowest section, could not attempt. This was contrary to the scheme of the work and unfortunate, as it unduly reduced the marks in those sections. There is no doubt, however, from the general trend of the numbers, except in the lowest section, that considerable improvement was taking place.

It now remains to present the results of the work in arithmetical reasoning of the two 'groups' after they had been put together again and worked a further test.

Table IV., showing the marks for arithmetical reasoning of the practised and non-practised groups compared in corresponding sections.

Marks for Three Preliminary Exercises in Arithmetical Reasoning.	Non-Practised Group A.			Practised Group B.		
	No of Children	Av. Mark Per Child in Preliminary Exercises	Av. Mark Final Exercise.	No of Children.	Av. Mark Per Child in Preliminary Exercises	Av. Mark Final Exercise
30 and over.	1	33.0	15.0	1	32.0	16.0
25 to 30	5	27.4	14.6	4	28.7	13.5
20 to 25	2	20.5	7.5	3	22.0	7.0
15 to 20	4	18.0	9.2	4	18.0	9.2
10 to 15	4	14.2	6.0	4	14.0	4.2

Taking the groups as wholes, we find that Group A scores an average mark of 20.6 (mean variation 5.8) in the preliminary tests, and 10.2 (mean variation 3.6) in the final test; whilst Group B scores 20.6 (mean variation 5.9) and 9.1 (mean variation 4.7).

It is difficult to resist the conclusion that the group practised in accuracy of computation shows no improvement over the group not thus practised, i. e., there appears to be no transfer of improvement in accuracy from arithmetical computation to arithmetical reasoning. That is the main problem; but a secondary issue arises as to the value of improvement in numerical accuracy in rule sums. Does this improvement transfer to the numerical accuracy with which 'problems' are worked?

Both sets of problems, the preliminary and the final tests, were marked for numerical accuracy. In this school, the

only units of accuracy counted in this marking were those which were found in steps necessary to the solution of the problem. This is not satisfactory scientifically, but it gives the result a practical value, for teachers wish to know whether the practised children can therefrom be expected to get more of their problems right. The results follow:

Table V., showing the numerical accuracy of groups A and B compared in the working of the preliminary and final problems.

Marks for Numerical Accuracy in Preliminary Problems	Non-Practised Group A			Practised Group B.		
	No of Children	Av Mark for Numerical Accuracy in Three Sets of Preliminary Problems	Av Mark for Numerical Accuracy in Final Set of Problems	No of Children.	Av. Mark for Numerical Accuracy in Three Sets of Preliminary Problems	Av. Mark for Numerical Accuracy in Final Set of Problems
Over 180	2	189.0	108.5	2	187.0	124.5
170 to 180	2	173.5	105.5	0		
160 to 170	2	164.0	73.5	0		
150 to 160	3	158.0	80.0	6	155.2	107.0
120 to 150	2	131.0	67.0	4	138.0	101.0
Below 120	4	105.5	71.0	4	92.7	*49.5

It is not fair to conclude from this experiment that no amount of improvement in numerical accuracy would transfer to accuracy in numerical reasoning, but only that the improvement shown in this case was insufficient to transfer. There is a presumption that it might not, since there is a clear transfer of improvement in numerical accuracy itself from rule sums to problems. I should myself have been prepared to find a slight improvement in the arithmetical reasoning, not perhaps from direct transfer, but from the fact that, the numerical work of the problems having become more accurate and mechanized in the one group than the other, a

*The children in this group, F. D., R. D., C. B., and A. H., showed no continuous improvement in the practice exercises in numerical accuracy; the continued practice at short intervals seems to have diminished their arithmetical capacity, which is not altogether an unexpected result with very weak pupils. In the other groups there seems to be a clear gain in numerical accuracy in the problematic work.

larger share of the pupil's available energy could have been given to the method of solving the problems, i. e., to the arithmetical reasoning involved. But this does not appear to have been the case.

V. SECOND SERIES OF EXPERIMENTS. SCHOOL "B."

A second series of experiments was carried out in a municipal girls' school situated in a very poor neighborhood. They were done with a Standard II. class of a high average age, namely, 10 years. The school was one in which numerical accuracy had long been a marked feature of the work, whilst the problematic aspect of arithmetic had received somewhat less attention. It will be remembered that the general plan of the experiments requires the tests and exercises to be taken in schools in which the conceptions of good arithmetical work differ.

A brief chronology of the experiment follows:

Tuesday, April 20th, 1909.

10.00 A. M. to 10.20 A. M. All the class worked arithmetical problems.

10.20 A. M. to 10.35 A. M. All the class worked rule sums.

Thursday, April 22nd.

10.00 A. M. to 10.15 A. M. All the class worked rule sums.

10.15 A. M. to 10.35 A. M. All the class worked arithmetical problems.

Friday, April 23rd.

10.00 A. M. to 10.20 A. M. All the class worked arithmetical problems.

10.20 A. M. to 10.35 A. M. All the class worked rule sums.

Tuesday, April 27th.

10.00 A. M. to 10.15 A. M. All the class worked rule sums.

10.15 A. M. to 10.35 A. M. All the class worked arithmetical problems.

On the results of the work in arithmetical problems, or, rather, on the number of correct rational processes in the

working of the problems, the whole class was divided into two equal groups, and from 9.50 A. M. to 10.30 A. M. on Wednesday, April 28th, Thursday, April 29th, Friday, April 30th, Tuesday, May 4th, Thursday, May 6th, and Friday, May 7th, Group B—hereafter called the practised group—worked exercises in rule sums for the purpose of producing improvement in numerical accuracy; whilst Group A—hereafter called the non-practised group—did exercises in drawing.

Next week the two groups were put together again, and, from 10.00 A. M. to 10.20 A. M. on Tuesday, May 11th, worked a test in problematic arithmetic. A second test was given at the same time on Thursday, May 13th; a third on Friday, May 14th, and a fourth on Tuesday, May 18th. After the preliminary tests, throughout the practice period, and during the final tests, no arithmetic was done by either group except the rule sums worked by Group B on the dates given.

i. *Specimens of Tests and Exercises.*

(a) *A Set of Preliminary Problems.*

- (1). How many cakes at 7 for 6d. can I buy with 14s. 6d.?
- (2). A man had £5. He gave a quarter of it to his son and 19s. 6d. to his daughter. How much had he left for himself?
- (3). What shall I pay for three pounds of tea at 1s. 8d. a pound, and two pounds of coffee at 1s. 3d. a pound?
- (4). In a wood there were 125 oak trees, 78 beech trees, half as many elm trees as there were beech trees, and 30 less ash trees than oak trees. How many trees were there altogether?

(b). *A Set of 'Rule' Sums worked side by side with a Set of*

Preliminary Problems.

- (1). Divide £1 19s. 11½d. by 12.
- (2). Find 8 times £3 13s. 4d.
- (3). From £13 6s. 11½d. take £9 19s. 3¾d.
- (4). £1 16s. 11½d. + £2 7s. 6d. + 15s. 9¾d. + £1 0s. 10½d. + £3 13s. 11½d.

The rule sums employed for practice exercises were similar, but more in number. The problem papers set in the final tests were intended to be of the same difficulty as those set in the preliminary tests, but they were really rather more difficult.

ii. *Method of marking.*

As before, the marks for reasoning depended on the logical steps necessary to the solution of the problem. Thus, in the case of the first sum in the set of problems given above, it is necessary to find how many sixpences there are in 14s. 6d., and then to multiply this number by 7. In the second problem, it is necessary to divide £5 by 4, to add the quotient to the daughter's share and subtract the sum from £5. This gives us 3 steps or marks. Another method, which subtracts the quotient from £5 and subsequently subtracts 19s. 6d. from this remainder, also requires three logical 'steps' or processes and equally receives 3 marks.

The marking for numerical accuracy in rule sums proceeded on the same basis as before—one mark was given for every accurate unit of addition, subtraction, multiplication or division.

In marking the numerical accuracy of the final set of problematic tests in this school a mark was given for every correct numerical unit whether it occurred in work necessary to the solution or not.

iii. *Results.*

I propose first to show a general table of correlation between accuracy of arithmetical reasoning and accuracy of numerical computation. The tests on which the figures are based, it will be remembered, were worked side by side, under the same school conditions.

Table VI., showing general correlation between
 (a) *accuracy of arithmetical reasoning and*
 (b) *accuracy of numerical computation.*

Marks in 4 Preliminary Tests in Arithmetical Reasoning	No of Children.	Average Mark for Reasoning.	Average Mark for Numerical Accuracy.
Over 35	4	36.5	198.0
35 to 30	5	32.6	195.2
30 to 25	8	28.8	184.5
25 to 20	7	23.2	184.1
20 to 15	9	18.1	148.6
15 to 10	2	12.0	121.5
10 to 5	7	9.0	94.3
5 to 0	1	5.0	39.0

It is obvious from the above table that the positive correlation is considerable, but it was thought advisable to work it out from the 43 individual cases by the Pearson product—moment formula. It reaches the high figure of + .79.

I am unable to show the improvement in numerical accuracy made by the practised group during the progress of the practice exercises themselves. Both the Head Mistress of the school and the Teacher of the class assert that it was considerable. Unfortunately the sums were only marked 'right' and 'wrong' and the children's worked papers were not preserved.

But that some improvement in numerical accuracy occurred and, moreover, transferred itself to the numerical accuracy of the final problematic work is obvious from the following table:

Table VII., showing numerical accuracy of Groups A and B
 (a) *in the preliminary problem tests,*
 (b) *in the final problem tests.*

Marks for Numerical Accuracy in 4 Pre- liminary Problem Tests.	Non-Practised Group A.				Practised Group B.	
	No of Children	Av. Mark 4 Prelimi- nary Tests.	Av. Mark 4 Final Tests.	No of Children.	Av. Mark 4 Prelimi- nary Tests.	Av. Mark 4 Final Tests
Over 150	7	171	139	7	163	142
150 to 100	8	129	110	8	120	117
100 to 50	4	82	76	4	76	74
50 to 0	3	42	67	3	37	74

It appears from the above table that, whereas the numerical accuracy of the non-practised group was decidedly superior to that of the practised group in the preliminary tests (the *equality* of the groups was based on their arithmetical reasoning, it will be remembered), yet, notwithstanding this, after the practice period, the position of the two groups, section by section, is reversed. Nearly all the sections appear to have gone down; a study of the possible numerical units and of the problematic difficulty in the preliminary and final tests show this to be appearance only. But even if it were so, we are concerned solely with the comparison of the final work of the two groups, and it is evident from these comparisons that Group B is the better. If we calculate the final results of each section of each group as a percentage of its preliminary result, we find the highest section of Group B scored 87%, whilst the corresponding section of Group A scored 81%; the second section of Group B scored 87% against the 85% of the second section of Group A; the third section of Group B scored 97% as compared with 93% for the third section of Group A; whilst the lowest section of Group B scored 200% as against the score of 159% for the lowest section of Group A. But it is not quite satisfactory to estimate improvement on numerical accuracy in this way, unless we can be fairly certain that the two groups would *attempt* an equal number of problems.

The more important question—the issue to which the experiment was mainly directed — was not, however, whether numerical accuracy was transferred from rule sums to problems, but whether improvement in numerical accuracy was transferred to accuracy in arithmetical reasoning.

The following table helps us towards a solution:

Table VIII., showing accuracy of arithmetical reasoning of Groups A and B in
(a) the preliminary problems. (b) the final problems.

Marks for 4 Preliminary Tests.	Non-Practised Group A.			Practised Group B.		
	No. of Children.	Av. Mark Prelimi- nary Tests.	Av. Mark Final Tests.	No. of Children.	Av. Mark Prelimi- nary Tests.	Av. Mark Final Tests.
Over 35	2	37.0	33.5	2	36.0	35.0
35 to 30	2	32.5	32.0	3	32.7	32.7
30 to 25	5	29.0	28.8	4	28.7	28.0
25 to 20	3	24.0	23.3	4	*22.5	27.7
20 to 15	5	18.6	18.8	4	17.5	16.0
15 to 5	5	9.0	12.2	5	9.8	15.4

*'E. F.', one of the children in this section, had been absent from school for some time previous to the preliminary tests and consequently worked these below her usual form; she was, according to the Head Mistress, improperly placed in the grouped and ordered lists.

Considering the groups as wholes, we find that, whilst Group A obtains an average mark of 22.5, with a mean variation of 7.9 in the preliminary tests and 22.7 in the final tests with a mean variation of 7.7, Group B obtains 22.5 (mean variation 7.7) and 24.2 (mean variation 7.4) in the preliminary and final tests respectively.

In this school the practised group and the non-practised group, which each scored 494 units of correct reasoning in the preliminary problem tests, are, as in the preceding school, *not* alike in the final tests, for now the former scores 532 units as against 499 for the latter. This is a reverse relationship to that found in School 'S.' It seems to me, however, that it would be unsafe to conclude that any transfer of accuracy had taken place unless, section by section, the children of corresponding ability throughout the two groups showed clear and consistent gains on the part of the practised group. In this case two of the non-practised sections are superior to the practised in the final tests.

VI. THIRD SERIES OF EXPERIMENTS. SCHOOL "I."

A third series of experiments was carried out in a municipal girls' school differing in type from the two previous

ones. It was situated in a rather good neighbourhood, but the arithmetical work of the children, both with respect to their capacity to solve problems and their numerical accuracy, was known and acknowledged by the teachers of the school to be low. Psychologists will see at once how important it is, in work of this kind, in which improvement is involved, to select observations from those who are at very different stages with respect to their *saturation point* for the mental operations in question. From a comparison with children of corresponding social class in other schools, I was well aware that these children were below saturation point in both accuracy of arithmetical reasoning and accuracy of numerical computation. The following is a brief chronology of the experiment:

On Tuesday, April 27th, from 9.40 a. m. to 10.0 a. m., the whole of a Standard III. class, of an average age of 10 years 4 months worked a test in arithmetical problems, and from 10.0 a. m. to 10.20 a. m. worked an exercise in rule sums. On Thursday, April 29th, the class worked further tests in problems and rule sums at the same times as before. On Friday, April 30th, further tests of similar character were worked at corresponding times. Each child had now worked three sets of problematic tests and three sets of rule sums. The class was then divided into two equal groups on the basis of the marks for reasoning obtained in the working of the problem papers; and on Tuesday, May 4th, from 9.40 a. m. to 10.20 a. m. one of the groups—hereafter called the practised group—worked ‘rule’ sums; the other group did an exercise in drawing. On Thursday, May 6th, at the same time in the morning, the practised group worked a second exercise in ‘rule’ sums; the non-practised group studied geography. Again, on Friday, May 7th, on Tuesday, May 11th, on Thursday, May 13th, and on Friday, May 14th, at the same time in the morning as before, the practised group worked further exercises in rule sums whilst the non-practised group studied history or geography, or were exercised in drawing.

Next week, on Tuesday, May 18th, the two groups were put together again, and, from 9.40 a. m. to 10.0 a. m., all the class worked a test in problematic arithmetic. A second test was given at the same time on Thursday, May 20th, and a third test on Friday, May 21st, also from 9.40 a. m. to 10.0 a. m.

i. *Specimen Sets of Tests and Exercises.*

(a) *A Set of Problems in Arithmetic.*

1. A man bought a chair for 30 shillings. He paid 9s. 3d. for a new cushion and 1s. 6d. for having it put on. He then sold the chair for £3 2s. 0d.; how much did he gain?
2. In a tramcar there were 29 people. At the first stop 8 got out and 5 got in; at the second stop 13 got out and 10 got in. How many were in the car then?
3. In a wood there were 972 trees, half were oak and a quarter elm, whilst the remainder were ash. How many ash trees were there?

(b) *A set of Rule Sums, worked during the Preliminary Tests by Both Groups.*

1.	£.	s.	d.	2.	$3,649 \times 79$
	35	10	$2\frac{3}{4}$		
	341	12	9		
	9	5	$11\frac{1}{4}$		
	38	19	$4\frac{1}{2}$	3.	$£50\ 17s.\ 11\frac{1}{4}d. \div 49.$
	264	4	3		

ii. *Method of Marking.*

As before, in marking the problems for reasoning, one mark was given for every rational 'step' towards the correct result, thus, in problem 1, it is necessary (a) to add expenses and (b) to subtract them from the selling price. Two marks are obtained if these processes are shown, whatever the result may be in numerical accuracy. In the case of problem 2, there are several modes of procedure, but each will be found to involve four logical steps or processes; thus, we may subtract 5 from 8, and 10 from 13, add the two remainders and then subtract from 29. Or we may subtract 8 from 29, add 5 to the remainder, take away 13 from the sum, and then add 10 to the remainder.

In marking for numerical accuracy, one mark as before was allowed for every correct process in addition, subtraction, multiplication or division.

iii. *Results.*

Table IX., showing general correlation in the preliminary exercises between

- (a) *reasoning as applied to arithmetical problems,*
(b) *numerical accuracy in the working of 'rule' sums.*

Marks for Reasoning in 3 Preliminary Exercises.	No. of Children	Av. Mark Per Child for Reasoning	Av. Mark Per Child for Numerical Computation.*
19	9	19.0	17.2
18, 17	8	17.2	17.7
16, 15	8	15.6	15.5
14, 13, 12, 11	6	12.5	15.8
7, 6, 5, 4	7	5.4	9.1

*The marks for each child in numerical computation were, in this school, tabulated to the nearest ten

It is fairly obvious that positive correlation exists, but it is also obvious that no very definite value can be obtained from the grouped averages. An application of the Pearson formula to the individual cases shows the positive correlation to amount to +.69.

Table X., showing the improvement of numerical accuracy within the practice-medium itself.

Marks for Reasoning in 3 Preliminary Tests	No. of Children	Av. Mark Per Child 1st and 2nd Practice Exercise	Av. Mark Per Child 5th and 6th Practice Exercise
19	5	123	163
18, 17, 16	6	122	149
15, 14, 13, 12, 11	5	119	123
6, 5, 4	3	63	87

There can be no doubt that, on the whole, a very considerable improvement (approximately 21%) in numerical accuracy has taken place. Let us now see how far (if at all) this improvement in accuracy is transferred to accuracy in arithmetical reasoning.

Table XI., showing Groups A and B compared in
 (a) *reasoning in preliminary problems,*
 (b) *reasoning in final problems.*

Marks for 3 Preliminary Tests.	Non-Practised Group A			Practised Group B		
	No. of Children	Av. Mark in 3 Preliminary Tests	Av. Mark in 3 Final Tests.	No. of Children.	Av. Mark in 3 Preliminary Tests.	Av. Mark in 3 Final Tests.
19	4	6.3	7.5	5	6.3	7.1
18, 17, 16	7	5.6	6.4	6	5.5	7.7
15, 14, 13, 12, 11	4	4.5	4.8	5	4.3	5.2
6, 5, 4	4	1.9	1.2	3	1.6	1.6

Taking the groups as wholes we find that whilst Group A scored an average mark of 14.3 (mean variation 3.9) in the preliminary tests and 15.7 (mean variation 6.0) in the final tests, Group B scored in the corresponding tests 14.5 (mean variation 3.8) and 17.9 (mean variation 6.2).

For our present purposes the most useful comparison lies between the marks in the three final tests for the corresponding sections of the practised and non-practised groups. There certainly appears to be an advantage on the side of the practised group in the 2nd, 3rd, and the lowest sections, but it is somewhat disconcerting to see that the highest section of the practised group is worse than the corresponding section of the non-practised group.

VII. FOURTH SERIES OF EXPERIMENTS. SCHOOL "O. K."

A fourth series of tests and exercises was undertaken in a municipal boys' school, situated in a rather poor neighbourhood.

The arithmetical work, both on its computational and problematical sides, was decidedly good. The experiment was carried out with the whole of a Standard III. Class, which had an average age of 10 yrs. 0 months, at the time of the commencement of the tests.

The following is a brief chronology of the course of the experiment:—

On Thursday, April 14th, 1910, from 9.55 a. m. to 10.25 a. m., the whole class worked a test in problematic arithmetic; and

at the same time in the morning a second, third and fourth test on April 19th, April 21st, and April 26th respectively. The class was then divided into two equal groups on the basis of the marks obtained for reasoning in the above tests, and, on May 4th, 5th, 6th, 10th, 11th, 12th, 13th, 18th, 19th, and 24th, practise exercises in rule sums were given to one of the groups for half-an-hour each day, whilst the other group worked exercises in English composition. During this period and until the end of the experiment, no arithmetic, other than the exercises and tests of the experiment itself, was done by either group. The irregularity in the succession of the practice exercises was due to intervals of school holiday. After the ten practice exercises, the two groups were put together again and worked four final tests in problematic arithmetic. The first was done from 9.55 a. m. to 10.25 a. m. on Thursday, May 26th; the second, third and fourth at the same time in the mornings of May 31st, June 2nd, and June 7th, respectively.

i. *Specimen Tests and Exercises.*

(a) *A Test in Problematic Arithmetic.*

1. At an election, one man has 6,913 votes, another 7,948, and a third 5,427. What was the number of votes altogether?
2. I had £5 in my pocket. I gave away 18s. 4½d. to one man, 10s. 6¾d. to another, and 13s. 4d. to another. How much have I left?
3. Tom has £8, and Harry has £6 12s. 6d. I give each of them as much again as he already has. How much more will Tom have than Harry after that?
4. A farmer buys 18 lambs at 12s. each. Six of them stray away and are lost. Those which remain he sells at £1 1s. each. What does he gain?

ii. *Method of Marking.*

The problems were marked as before, without reference to the numerical accuracy of the result, and solely with reference to accuracy of method. It is obvious that Problem 1. is a 'one-step' problem; the only rational process is the

judgment that the numbers must be added. Problem 2. involves two processes, the addition of the sums given away and the subtraction of the total from the original sum. Problem 3. is a 'three-step' sum to children at this mental level. Tom's money must be doubled, Harry's money must be doubled, and one of the sums thus obtained must be subtracted from the other. Problem 4. is a 'four-step' problem. It is necessary to find the cost of the lambs, to subtract the number which stray away from the number bought, to find the amount realized by the sale of the remainder, and to subtract the cost price of the lambs from the selling price.

The practice exercises in rule sums were marked for numerical accuracy in the same way as in previous experiments.

iii. *Results.*

It will doubtless have been noticed that in this experiment no tests in numerical accuracy were set *pari passu* with those in accuracy of arithmetical reasoning. Any correlations, therefore, which are worked out between these two sets of operations can be found only for that half of the class which took the subsequent practice exercises. I do not consider this is so satisfactory theoretically, since the operations are not made under precisely the same conditions. But a correlation worked out on the basis of *four* separate sets of tests for reasoning and *ten* separate exercises in numerical accuracy cannot fail to have some value.

Table XII., showing the correlation between
(a) accuracy in arithmetical reasoning,
(b) accuracy in numerical computation,
of the Practised Group only.

Marks for 4 Preliminary Tests in Reasoning.	No. of Boys.	Av. Mark Per Boy in Reasoning.	Av. Mark Per Boy in Numerical Computation.
35 to 40	4	37.8	1916.7
30 to 35	3	34.0	1823.7
25 to 30	5	28.0	1573.0
20 to 25	5	22.6	1309.2
15 to 20	8	17.5	1285.0
10 to 15	4	12.5	1222.0
5 to 10	4	7.8	1057.5

The correlation is obviously a high one, and its exact determination from the individual cases by means of the Pearson formula of correlation gives $+.85$.

I will now show the improvement in the practice medium itself. The classification, as in the above table, is based upon the marks obtained in the preliminary tests in arithmetical reasoning.

Table XIII., showing the improvement in numerical accuracy within the practised group.

Marks for A Preliminary Tests in Reasoning.	No of Boys.	Marks for 1st and 2nd Practice Exercise.	Marks for 9th and 10th Practice Exercises.	Percentage of Improvement.
35 to 40	4	323	410	27%
30 to 35	3	316	397	26%
25 to 30	5	278	329	18%
20 to 25	5	270	313	16%
15 to 20	8	250	310	24%
10 to 15	4	217	253	17%
5 to 10	4	192	230	19%

It will be seen that the improvement is, in all cases, very considerable; it now remains to be seen how far this improvement in accuracy of numerical computation is transferred, if at all, to accuracy in arithmetical reasoning.

Table XIV., showing the comparison between the practised and unpractised groups in

- (a) *four preliminary tests in arithmetical reasoning,*
- (b) *four final tests in arithmetical reasoning.*

Marks for 4 Preliminary Tests in Reasoning.	Non-Practised Group A.			Practised Group* B.		
	No. of Boys	Av. Mark for 4 Preliminary Tests.	Av. Mark for 4 Final Tests.	No. of Boys.	Av. Mark for 4 Preliminary Tests.	Av. Mark for 4 Final Tests.
35 to 40	4	37.8	38.3	4	37.8	38.8
30 to 35	3	34.0	38.7	3	34.0	37.0
25 to 30	6	28.0	28.2	5	28.0	32.8
20 to 25	4	22.8	23.8	5	22.6	28.6
15 to 20	8	17.8	19.5	8	17.5	20.3
10 to 15	5	13.2	18.8	4	12.0	13.8
5 to 10	5	8.4	8.6	4	7.8	13.3

*Two boys, E. W. and C. H., are omitted from the practised group. as one was absent on 8 occasions and the other on 3 occasions during the practice exercises.

Considering the groups as wholes, we find that Group A obtains an average score of 21.8 (mean variation 8.1) and 23.6 (mean variation 9.2) whilst Group B obtains 22.0 (mean variation 8.1) and 25.5 (mean variation 8.5) in the preliminary and final tests respectively.

It is obvious that, on the whole, the practised group does better work in the final tests; in the first of them this group obtains an average mark of 6.2, in the second of 5.7, in the third of 7.6, and in the fourth of 6.0; whilst the corresponding marks for the non-practised group are 5.6, 5.2, 7.1 and 5.7. But the excess improvement of B over A, section by section, is uncertain; indeed, in two cases, the non-practised section does better work than the corresponding section of the practised group.

Every section in both groups shows improvement in arithmetical reasoning, though no arithmetical exercises other than the rule sums, and those by the practised group only, were done during the experiments. This general improvement is probably due to natural growth, and perhaps the influence of the other branches of school instruction; but this latter point is doubtful and needs proof. The general improvement is more obvious in this school because very great care was exercised in order to make the final tests in reasoning of no more intrinsic difficulty than the preliminary ones.

But an improvement which is common to both groups is not an improvement on which we can base the doctrine of the transfer of accuracy. To possess irrefragable support for a doctrine of that kind we should require an excess improvement, section by section, of Group B over Group A. Let us calculate the percentage of improvement from the preliminary to the final tests in each section of each group and see how far such an excess improvement is to be found.

Table XV., showing the percentage improvement (to the nearest unit) of each group, section by section, in arithmetical reasoning.

Marks for 4 Preliminary Tests in Reasoning.	Non-Practised Group.		Practised Group.		Percentage Gain or Loss of B Over A.
	No. of Boys.	Percentage Improvement.	No. of Boys.	Percentage Improvement.	
35 to 40	4	1%	4	2%	Gain 1%
30 to 35	3	14%	3	9%	Loss 5%
25 to 30	6	1%	6	17%	Gain 16%
20 to 25	4	4%	4	9%	Gain 5%
15 to 20	8	9%	8	16%	Gain 7%
10 to 15	5	42%	4	15%	Loss 27%
5 to 10	5	2%	4	70%	Gain 68%

VIII. GENERAL CONCLUSIONS.

My first conclusion is that much more of this work needs to be done; I hesitate to draw hard and fast conclusions, even negative ones, where there is so much irregularity in the results.

My second relates to the inferences which we may draw from high positive correlations between two quantities. It seems to be possible to find highly correlated functions which appear to have very little relationship of pedagogical value. We cannot conclude without further inquiry on other lines, that two highly correlated mental powers are causally related. If they have a common factor or a common cause, it may be one which our methods cannot influence, and its determination has then little value for practical direction, except in a negative sense. Two quantities may be highly correlated but the ratios of their growth may not be; nor may we be able to produce increase in the one by producing increase in the other. I understand that these empirical conclusions are quite in accord with correlational theory, but it is well to state them here, as it is quite possible that educationists will regard an established high correlation as justifying several inferences which do not properly flow from it.

My third conclusion is more especially pedagogical. Teachers wish to know whether they can by improving the

accuracy of function 'A' thereby influence for good the accuracy of function 'B.' I think that in some cases they can. I believe, for example, that various mnemonic functions are so connected in children and young people that improvement 'transfers.' We then need to know how much, for we should not be wise to set out to improve any function indirectly unless an indirect method were quantitatively a better method than the direct one; unless of course, the direct method were not available.

But as to accuracy of arithmetical reasoning and accuracy of numerical computation, what can we say?

It seems possible to improve the accuracy of numerical computation *without any certainty* that we shall thereby improve the accuracy of arithmetical reasoning.

For the present, therefore, pending more conclusive experiments, numerical accuracy should be sought for because it is valuable in actual life, and not because we can feel confident that an improvement in it will transfer to accuracy of arithmetical reasoning.

But if improvement in accuracy of numerical computation is not transferred, how shall we account for the general trend of the final results in reasoning (in all schools except 'S') in favour of the practised groups?

There seem to me two possibilities. I have judged the children's reasoning powers in arithmetic by means of problems numerically worked out, though the accuracy or inaccuracy of the numerical solution has not affected the mark for reasoning. Might not the greater facility in numerical computation, such as was obtained by most of the practised groups, set free, as it were, more mental energy to deal with the rational solution of the problems?

Also, since, in school, the functions of numerical computation and of arithmetical reasoning are often exercised together, the continuance of numerical computation by one of the groups and its discontinuance by the other might well maintain in one case a greater readiness for the associated

operations than in the other, and the former would on that account do better work in arithmetical reasoning than the latter.

But if either of these hypotheses were true, ought we not to expect more regular results; though not perhaps such regular results as if there were a direct transfer of accuracy?

It seems fitting to conclude this research with a question.

The value of the foregoing investigations appears to me to lie chiefly in the suggestion of a method by which a very important pedagogical question may be ultimately solved under school conditions. It is a method which teachers can easily use for themselves, and publish the results for the guidance of their colleagues. I am strongly convinced that, without such active co-operation on the part of the teacher, experimental pedagogy will remain an academic study, with little effect upon the practice of the schools.