EXPERIMENTAL RESEARCHES ON LEARNING TO SPELL.

BY W. H. WINCH.

London, England.

(Concluded from the November Number.)

Part II.

V. A SECOND EXPERIMENT, WITH STANDARD IV. BOYS.

Average Age 11 years 0 months.

i. *The Boys who did the Work, and why it was done with older Boys.*

In the first experiment we found that, with boys who were young and mentally undeveloped, a method of learning spelling, which, with attention directed by the experimenter, possessed marked articulatory and auditory factors as well as visual ones, was superior in results to a method in which the boys worked silently by themselves, and therefore apprehended the sequence of the letters mainly by visual means. I say *mainly* by visual means, since, as a matter of fact, the boys also muttered the words and letters as they were learning them. With young children of an undeveloped type it is practically impossible to prevent this. So that it might fairly be argued that by *both* methods the children really employed articulatory and auditory factors, as well as visual ones. But older boys can really work silently, or at least, much more so. Perhaps *really* silent visual work would be more efficacious than simultaneous work by a combined method. To test this hypothesis a similar experiment was carried out with the boys of Standard IV¹—the highest class in this school which contains, as an American would say, the primary, but not the grammar grades. As before, the boys who did the work were exclusively the children of Jewish aliens.

¹This Standard corresponds to Grade V in America and to the Fifth School Year in Germany.

ii. *The Lists of Words of Equal Difficulty.*

Two sets of words were selected, 30 in each set, from a small dictionary called '*Green's Dictionary of Selected Words, being the words added to a Scholar's Vocabulary in the Upper Standards of an Elementary School.*' The title is unduly hopeful, and I am not prepared to advocate that every one of the words in these lists should be known by all elementary pupils; but that non-advocacy will in no way invalidate the conclusions of our experiments. The following words were picked out:

The List of Words called Set A.

aborigines	exchequer
acrobat	extinguish
advertisement	disease
aeronaut	eclipse
alibi	committee
amateur	conciliate
analyze	congestion
apostrophe	conscience
apprehend	convulsion
arsenal	corroborate
asphalt	denominator
bachelor	destitute
blaspheme	dispensary
campaign	duplicate
chronometer	effervesce

The List of Words called Set B.

accomplice	excavate
acquit	extravagant
apprentice	eccentric
aggravate	economy
almanac	complexion
ambassador	conceal
annuity	conjuror
apparition	convalescent
archipelago	correspond
artificer	crucible
auxiliary	deteriorate
bazaar	dislocate
bronchitis	discipline
carnivorous	ecclesiastic
chrysalis	encyclopedia

I print these lists, since they have been tested and found equal, for one class at least; and they may be useful to other experimenters.

iii. *How the Test Words were Administered.*

They were called out by the experimenter—the Head Master, as before—each word once, and then written down, one by one, in small hand, by the boys taking part in the experiment. Words from Set A and Set B were called out alternately, one from Set A and then one from Set B, and so on. They were dictated at 2.15 p. m., after the second Registration on Monday, February 17th, the exercise taking 25 minutes. For the present I ask the reader to accept the equality in difficulty of the two sets of words—the detailed results will be given later.

iv. *The Chronology of the Teaching Practice for Set A and Set B.*

Each lesson lasted 10 minutes, and commenced either at 9.45 in the mornings, after Scripture lesson, or at 2.05 in the afternoons, after the first Registration. The words of Set A were learnt on the Tuesday morning following the preliminary test, on Wednesday afternoon, and on Thursday morning. Those of Set B were learnt on Tuesday afternoon, on Wednesday morning, and on Thursday afternoon. Each set of words, therefore, received three practice periods of 10 minutes each. Set A probably has a slight advantage in point of time for work of this kind; though, for work employing what are usually called the 'higher mental functions,' the early afternoon is probably better. But if any advantage exists, it is certainly a very slight one.

v. *The Two Methods Employed, one for Set A, the other for Set B.*

The method for Set A was precisely the same as that employed for Set A in the previous experiment in the same school; it was a combined method with directed attention. For Set B the method was, as before, a visual method without directed attention. All the testing and teaching were done by the same man throughout—the Head Master of the school. This is important with experimental work with children, since, to alter your experimenter as well as your method, is to run grave risks that a new factor, producing variation, may come in, to wit, the disciplinary influence of the experimenter. Sometimes this risk must be taken, but it should be avoided if possible. The reader is requested to refer to the account of the previous experiment for a psychological analysis of the two methods: these boys did not mutter; the method for Set B was, therefore, really visual, and visual only.

vi. *The Final Tests with Set A and Set B.*

On Monday, at 2.05 p. m., the two Lists of Words were dictated, word by word, one word from each set alternately. Set B, therefore, in the final testing, had the slight advantage of a more habitual time to set against the slight advantage of the larger number of *morning* lessons accredited to Set A.

vii. *The Results in (a) The Preliminary Tests, (b) The Final Tests.*

The complete results for each individual were tabulated and entered on a schedule. I suggest the general utility of such schedules in experimental pedagogy; they will save us from many hasty conclusions, and are of enormous interest to the teachers. Give a teacher such a schedule, with marks that he can understand, set out for the boys he knows (the psychologist doesn't, let us remember) and he will thank you and ask you to do some more work in his school. The teachers who know the boys will discuss the results most eagerly; they will tell you that you are wrong in certain cases—you probably are—but that will not prevent them (and they know, if anyone does, the variability of school-children) from frankly accepting the general results of the experiment, if they are set out in a form which they understand. The more sophisticated teachers are suspicious of mathematical calculations: the superior person, arguing from data by methods they do not know, will *never* convince *them*. I am not prepared wholly to condemn their attitude: I like to see individual results myself.

Table III, showing how the Individual Results were tabulated.

Name.	Tests before Learning		Tests after Learning	
	Number of Words Correct.	Number of Words Correct.	Number of Words Correct.	Number of Words Correct.
	Set A.	Set B.	Set A.	Set B.
M. Marks	15	15	28	26
A. Dorin.....	13	10	24	19
A. Roth.....	9	11	30	27
.....	..	..	..	..
P. Müller.....	7	6	24	24
.....	..	..	..	..
H. Ritts.....	3	3	23	22
S. Altman.....	3	3	24	19
.....	..	..	..	..
M. Serkuer.....	0	0	1	1
.....	..	..	..	..
H. Wolfstein.....	0	0	12	5
54 boys. Averages.....	4.3	4.4	17.9	15.1

Comments on Table III.

Apparently the words for Set A and Set B were well chosen: they are equally well done *before* learning. There has been enormous improvement *after* learning: and reference to the full schedule of results shows that there are only three cases

in which the improvement by the method employed for Set B is greater than that resulting from the method employed for Set A. A grouped correlation table follows.

Table IV., showing, section by section, the Number of Words Spelt Correctly before Learning and after Learning.

Number of Words of Set A Correct before learning.	No. of boys.	Before Learning.		After Learning.	
		Average Set A.	number right. Set B.	Average Set A.	number right. Set B.
Over 10.	3	13.0	12.0	26.3	24.3
9, 8.	8	8.5	8.5	24.5	22.7
7, 6.	8	6.9	7.1	22.4	18.4
5, 4, 3	14	4.0	4.6	21.4	16.4
2, 1.	10	1.6	2.6	12.3	11.5
0.	11	0.0	0.3	8.7	6.7

Comments on Table IV.

These are evidently not 'chance' results. All along the line, for the weaker pupils as well as for the abler ones, the simultaneous combined method with attention directed by the experimenter has shown itself victorious. To the statistician the result emerges quite as readily and more succinctly from the figures given in the following section:

viii. Statistical Evaluation of the Result.

The average number of words correctly spelt in Set A *before* learning was 4.3, for Set B it was 4.4. The correlation coefficient between the results of Set A and Set B was $+.81$, and the 'probable error' of the difference between the averages was .20. There is, therefore, no 'significant' difference between the two lists before learning.

After learning, the averages were, for Set A, 17.9, and for Set B, 15.1: the correlation-coefficient was $+.89$, and the 'probable error' of the difference between the averages was .34. The difference between the means, namely, 2.8, is 8 times its 'probable error,' and is therefore a highly 'significant' difference.

ix. Conclusions from the Second Experiment.

It was suggested that the presence of slight articulatory and auditory factors in the method used for Set B *might* have lessened somewhat the results by this method with the young children of Standard II. On the other hand it was also sug-

gested that a purely visual method might have a better chance with older children. So far, all that we can say is that with these older boys—about $2\frac{1}{3}$ years older—the result is in the same direction as with the younger children, but even more decided. The combined method has again won the victory. But, for there is a but, I am not yet convinced that for all elementary school children the combined method is the best. These boys are still young in years—about 11—and are certainly not of a highly developed type; and we must try the experiment in other schools of a different character. Still, for boys of this age and development, there is no doubt about the results; they are conclusive. But let us bear in mind that for Set B the attention of the boys was not directed, and for Set A it was. The difference may lie there, and not in the combination of motor, auditory and visual memories. But under the conditions given—a method actually in use in schools—the ‘combined’ method proves to be distinctly superior.

VI. SUBSIDIARY EXPERIMENTS WITH STANDARD I. BOYS.

Average age 7 years 0 months.

i. *How the Experiment came into Being.*

I call these exercises now to be described *subsidiary* experiments, though they were more numerous and extended over a much longer period than those which have previously been recounted. But they are really subsidiary in two respects. First of all, they preceded these experiments in time, and secondly, there were imperfections in them, considered experimentally, which render them less conclusive; though, as I have said, they were much more numerous and much more protracted. But collectively they have some value; and they are interesting as showing, in one actual case, how a principal, just appointed to a school, set about trying to improve the spelling; at first, with a fair degree of confidence. They also indicate how he began to realize that he did not really know the best way, and how he asked for guidance in finding out. The school was the ‘down East’ school in London in which, so far, all the spelling experiments described have taken place. The boys read fairly well, but it was said they spelt very badly.

The boys who did the work were graded as Standard I—a normal group of boys, since the Infants' department contained no class graded as a Standard class.

ii. *The First Method Employed to Improve the Spelling.*

Near the close of every afternoon session, six easy words, whose pronunciation and meaning were known to the boys, were written on the blackboard. The pupils were provided with slips of paper, and they were told to copy the words and learn them at home. The circumstances were not wholly unfavourable for home lessons; since, though the children lived in poor, and, in many cases, in crowded homes, their Jewish parents were keen—unlike so many of our English ones—that their children should learn and so get on in the world. The boys copied the words, but did they all learn them at home? Every morning the teacher of the class dictated the six words that the boys had been instructed to learn the evening previous. On the Monday following the week in which the 30 words had (or had not) been studied—six words had been given for each of five evenings, Monday to Friday inclusive—the Head Master dictated the whole list of 30 words. Teachers will be prone at first to think that a description of such a common experience is hardly worth while. But the value of it lies in the continuance of the method, and the careful tabulation of the results. For twelve weeks this process was repeated, from August 26th to November 15th. No improvement at all took place in the power to acquire fresh spellings in the class as a whole. Of the first 30 words, dictated on September 2nd, an average of 17.0 was spelled correctly; of the last 30 words, dictated on November 18th, the average spelt correctly was 17.0. Thirty per cent. of the class had shown no improvement any where along the series, and those who had improved had done so with many back-slidings.

The time taken averaged ten minutes a day, with an additional twenty minutes for the Head Master's weekly test, making a total of seventy minutes a week of school time for twelve weeks practice which showed practically no improvement. There certainly seemed no "transfer" improvement, the boys did *not* acquire the later sets more thoroughly and expe-

ditionously than the first. But of course they had learnt to spell, or many of them had, a number of words which they did not know before. The Head Master was not satisfied; he knew that, when learning by "good methods," children show steady and continuous improvement. In this case, after twelve weeks work, if they were improving at all, they were only creeping along, and some were not even creeping. And the school time: 70 minutes a week for it; was the result worth it? Wasn't the whole thing a futility? So he started out again by another method.

iii. *The Second Method Employed to Improve the Spelling.*

Hitherto, it was certain that the boys had copied the words, and had attempted next day to write them down when they were dictated, but whether they had made any visual or other repetitions of them at home was not known. Probably some had and some had not. The method had failed to produce rapid improvement, and it was costly in school time. A start was therefore made by another method on December 2nd. As before, the object was to teach, or require to be learnt, 30 words, whose spelling was, generally speaking, unknown, within one week. The lessons, ten minutes long, were taken at the commencement of every afternoon session, five lessons a week, at 2.05 p. m. The words selected may prove of interest to teachers. They were as follows:

bird	mouse	butterfly	flower	bluebell
sparrow	bear	snake	daisy	fruit
goose	wolf	worm	lily	orange
swan	monkey	wasp	wallflower	pear
gooseberry	carrot	sugar	coffee	wine
turnip	bean	water	cocoa	brother

The words were written on large cards (12 inches by 15 inches), one word on each card. The boys were provided with pens and slips of paper. Each word was shown to the class for 10 seconds. At the commencement of each exposure, the Head Master pronounced the word in a clear loud voice. At the end of the ten seconds the card was removed and the boys were told to write the word down quickly. Twenty seconds were allowed for writing, and the next word was pronounced as before. The boys were instructed to learn the words by looking

at them, but, in actual practice, they muttered the letters as well, whilst they were looking at them; but it was a subdued, and, as it were, a contrary-to-order muttering, which made it rather a surreptitious affair. The first ten of the thirty words were exposed, studied and written from memory, one at a time, within five minutes: in the remaining five minutes of the lesson they were repeated in exactly the same way as before. On the second day the next ten of the thirty words were studied in precisely the same way, and on the next day the third or concluding ten. On the fourth day, twenty skeleton words were placed on the blackboard, for example, ird: s arr w: g se: etc., and the boys were told to complete as many as they could. They liked the exercise, but found it difficult: very few boys completed the whole twenty. On the fifth day the ten remaining words were skeletonized, and also ten of the most difficult in the list. The boys did not work at the completion of them with much success. On the following Monday morning the whole list was dictated. There was an average number of correct spellings amounting to 17.7. "Not a very great advance on the previous method," said the Head Master. "I believe the 'skeleton' words were of little use and I shall discard them." The method was costly in time too, seventy minutes for 30 words. I shall try again."

But some of the weaker boys spelt better than they had ever spelt before.

iv. *The Third Method Employed to Improve the Spelling.*

Once again it was proposed to require 30 words to be learned within a week, to give 10 minutes a day to them, and, as before, to have the work done visually without overt articulatory and auditory factors.

The words selected were as follows:

round	feed	found	jumps	folk
line	uncle	thought	kitten	miss
own	twice	poor	fight	back
lamp	early	bricks	always	frock
pocket	catch	bark	arm	tart
awake	house	candle	night	rock

We did not know how many of these words the boys could spell already, before the special practice; and it is possible

that the boys were by growth and practice at a different slope of their curve, so to speak, this week than last; and these words also may be easier or harder than the last set. Teachers would, I think, say they were easier. It was the consideration of points like this which forced the Head Master to ask for help in framing more conclusive experiments. Still, if this work is corroborative of more exact experiments, it will not be unprofitable to record it.

On Monday, at 2.05 p. m., the first ten of these words were exposed, one at a time, and pronounced by the experimenter as in the second method, twenty seconds being allowed for each word to be written from memory by the boys. This took 5 minutes; the same words were given again in the same way during the remaining 5 minutes of the lesson. On Tuesday the next ten words were studied, and on Wednesday the third and concluding ten. Each word had, so far, been exposed twice and written from memory twice. On Thursday ten of the hardest words in the set were exposed and written twice; and on Friday—the fifth day—ten more of the harder words were dealt with. The twenty harder words had, therefore, been exposed and written from memory four times. There was, on this occasion, no ‘skeletonizing’ of the words. On the following Monday morning, December 16th, the whole list of 30 words was dictated. The average number correctly spelt was 20.8 and every boy in the class was able to spell *some* words correctly. This looked like a distinct improvement, but the method was costly in time; and was rather complicated. For there were, undoubtedly, some articulatory and auditory factors employed, as well as visual ones, for the boys ‘forgot themselves and muttered’: and there was a graphic motor factor as well, with a renewed visual stimulus working favourably for those who wrote the words *correctly* from memory after exposure, and unfavourably for those who did not.

v. *The Fourth Method Employed for Improving the Spelling.*

The teacher of the class, a young man fresh from Training College, now raised an objection to the Head Master’s procedure. He favoured a frankly open appeal by a combined visual, auditory and motor-articulatory stimulus. He

thought the Head Master's method was too predominantly visual and that he himself could do better. So the Head Master and he decided to teach for a week against each other. They picked out two lists of words of apparently equal difficulty,² and taught each by his own method. They did *not* test the equality of the lists beforehand, as they should have done; but they timed each other's work, and kept each other rigidly to the methods stipulated at the outset. The assistant took the mornings, at 9.45, the Head Master the afternoons, at 2.05. The Head Master taught his 30 words precisely as in the third method, 10 minutes for each of five lessons. The assistant teacher's method was frankly a combined one. On the first day ten words, written on the blackboard, were, one by one, spelt and pronounced once by the teacher and then spelt and pronounced, one by one, twice by the boys in simultaneous work. After that the ten words were spelt through again, twice each, by the boys, without the teacher's example. Then the hardest words of the ten were spelt once again. The weaker spellers then, without the other boys, spelt the words twice. The stronger spellers were now, with their backs to the blackboard, called upon to spell the words, whilst the weaker pupils checked them from the board. Then the weaker boys were also called upon to spell with their backs to the board.

This procedure seems a good deal to get through in ten minutes but the Head Master was checking the time, and it was religiously observed. On the second day the next ten words were dealt with in the same way; and on the third day, the

²List of Words Taught by the Head Master.

week	ship	soap	Friday	four
minute	whistle	wax	Summer	six
month	salt	candle	Winter	eight
time	iron	silk	holiday	ten
ball	clay	Monday	Christmas	second
hoop	lead	Wednesday	two	fourth

List of Words Taught by the Class Teacher.

day	doll	glass	Thursday	three
year	bat	gas	Spring	five
hour	coal	string	Autumn	seven
night	zinc	paper	season	nine
toy	stone	Sunday	Easter	first
train	gold	Tuesday	one	third

third and concluding ten of the thirty words. On the fourth day—Thursday—the boys spelt and pronounced all the thirty words twice, not after the teacher. This occupied some 4 minutes: the remaining six minutes were taken up by the boys writing down as many of the 30 words as could be done in the time from dictation. On the fifth day all the words were spelled again twice, by the boys, and again the rest of the lesson was given up to dictation.

I need hardly say that these little dictation exercises were not returned marked to the boys; it might have given an undue advantage to the class teacher's method, and would have been costly in time. On the following Monday the two lists were dictated. The average number of words spelt correctly by the Head Master's method was 22.9, and by the assistant's method was 23.0.

The Head Master's method was less laborious in teaching, for he did nothing but pronounce and expose the words; the assistant's method was much more laborious for him. The comparative results were close; each stuck to his own opinion, and in the conflict of argument called upon me to decide. I said that I did not know, but that I would show them how to find out which was the better method. Couldn't I tell, said they, from some mathematical or other analysis of the figures? The answer was in the negative. A consideration of the reasons given will show how we were forced to use the exactly controlled procedure of the First and Second Experiments.³ First of all, we did not *know* that the two sets of words just employed were equal in difficulty for the class. Secondly, the two methods were taught by two teachers *not* of equal experience and skill. Thirdly, the one lesson did not enjoy quite the same advantage in the placing of the lessons as the other, for one was in the mornings at 9.45, and the other in the afternoons at 2.05. And fourthly, I said, we should have to work with fresh and unsophisticated children; for these Standard I children were, by now, unequally practised; the Head Master's method was too dominant. Then the methods, if we wanted to *prove* anything, should be kept more distinct from each other.

All these conditions were observed; and the researches

³Part I. and Part II. of this Research. 1

directed by me, and called 'A First Experiment' and 'A Second Experiment,' followed.

vi. *Conclusions from the Subsidiary Experiments.*

The outstanding conclusion is of course the need for *controlled procedure with exact measurement*. Point by point, by the logic of events, teachers who try these issues for themselves, are driven along until they reach conditions of experimenting which they see for themselves are valid. They cannot usually invent valid experiments, but they can carry them out (with a year or two's practice); and they are admirable critics of what constitutes validity. They come to an *impasse*, and appeal for guidance. This guidance it is the duty of every educational authority to provide. I do not mean that the authority should appoint superior persons to *tell* the teachers what methods to adopt: the superior persons do not really know any more than the teachers do. I mean that *educational psychologists should be appointed who could help the teachers to find out*. Dozens—hundreds of teachers could set to work (under guidance) to *make the science of experimental pedagogy*. The educational psychologist wants no coercive powers: the teachers *will come to him*. And they will hold debates among themselves, each describing his own experiments, and criticising those of others. Mere talk will be at a discount, and the men and women who can *prove* good methods will come into their own. Is this a mere ideal? Oh no, it is what freedom of method to teachers and encouragement to them to experiment will inevitably lead to.

To return: from the *subsidiary experiments* the combined method seems likely to prove the better for these children, but the imperfections of procedure drive us on to the controlled procedures of the First and Second Experiments. The subsidiary experiments illustrate genetically the pathway to experimental pedagogy.

vii. SUMMARIZED CONCLUSIONS FROM PART I AND PART II OF THE RESEARCH.

The 'Subsidiary Experiments' preceded the 'First' and 'Second' Experiments in order of time. From them we were driven on to the more carefully controlled experiments known

as 'First' and 'Second.' As far as anything was indicated by the 'Subsidiary' experiments, they showed a balance in favour of methods with combined stimuli and directed attention. But the children who did the work were of low grade (Standard I) and very young (from 7 to 8 years of age), and it is a commonplace of pedagogy that young and mentally undeveloped children require external stimuli to direct and maintain their attention to an extent which their elders do not. With Standards II and IV, children ranging from 8 to 11 years of age in the same school, however, similar results were obtained. The method with pronounced auditory and articulatory factors, as well as visual ones, and with more direction of attention by the experimenter, showed itself definitely superior to a method more purely visual and with less direction of attention. I do not propose to generalize the results beyond the conditions of the experiment. This method may not be the better for English children, nor for any children of a superior mental type. I am prepared to accept the possibility that children, trained to a greater extent to work by themselves than these children had been, might do better on a visual silent method. I also stress the point that it is *not* recommended that lists of words should be picked out irrespective of their meaning, their orthographical construction, or their occurrence within a graded system of teaching reading. Moreover, we have not disproved the feasibility of success by an indirect or incidental method. But we *have* shown that when new spellings are to be acquired by children of this age and mental type, a method which contributes pronounced articulatory and auditory factors, as well as visual ones, is more successful than a method which depends almost wholly upon the latter.

The question of analytic or synthetic methods of teaching reading and orthography with *very* young children is not raised. This is Senior School work, and in Senior Schools the orthography to-day is almost invariably taught by an alphabetic (not phonic) spelling method of some kind, with visual, auditory, articulatory, graphic stimuli or combinations of these. To teachers in Senior Schools in poor neighbourhoods this research especially appeals.