TESTING AND EVALUATING

By

JOHN FLAVELL

YOU may be wondering why testing and evaluating has a place relatively early in this course, as it deals only indirectly with mathematics, but I am sure you will agree that the content and the nature of the testing that children undergo have a most profound influence on the nature of the teaching in the schools. We are all human enough to tend to teach what is going to be tested. Consequently, testing is of the first importance. If you want a change in your mathematics teaching, one of the best agents for that change is a modification in your testing procedure.

The great evil of testing is not so much the testing as the use made of the testing. Take, for instance, this business of "labelling" children. I am all for finding out as much as we possibly can about children, but that information should not be used as a label—labelling a child A or B or C, good in this or bad in that—that information should be private, should never be disclosed at all, except under special circumstances. I think the English-speaking world, particularly England and America, have in the past been completely test-mad. Is it not time we thought about this testing a little more seriously? Testing takes time. Testing is in the main totally negative. Is it not time we wondered whether the time we devote is well spent? Would we not do better to spend our time on something constructive, something positive, something adding to the child? So I would make my first point: consider what time you feel is justified in the testing situation.

The second point I would like to make is this: If teachers were people of integrity, they would pay little attention to an examination. I also feel that, rather than being a foresworn evil, if the examination is properly devised, it can be a force for good. It can be a healthy influence, but if it is going to be that, I would suggest the examination or the test should be drawn up by the teachers of the children who will take the test. I think a test should in most cases be home-made.

Now I know that raises certain difficulties. It usually costs more money. But in a city built on gold, I do not think that could be of great consequence.

In a testing situation we must be careful of relying only on examination results, the result of a test or battery of tests, because such tests are highly dubious in their reliability (and I include in that statement all the standardized tests you like with those dreadful things they call norms and mental ages). We have been bedevilled with those wretched things for years and years, and the only people, I think, who have gained anything from them are the statisticians. A statistician is like a drunken man who leans on a lamp-post for support rather than for illumination. Disraeli said: "There are lies, damned lies and statistics."

Objective tests are highly suspect. What can we use to supplement them? I would say that any series of tests in any kind of work are incomplete unless they are reinforced by the teacher's observation. I would even go so far as to say that you do not need to test children—ask the teachers, they will tell you. Any experienced teacher knows perfectly well which of his children are capable of benefiting from a certain type of education or other, but even there there are difficulties, the human element comes in again. I would like to see always some objective rating put alongside the subjective, but to me the subjective is far the more important.

I would like now to take a look at various aspects of the type of test which I would like to see in the primary schools—and I can speak only of primary because my experience has been almost entirely in primary schools. What do we test? Do we test facts, routines, skills, potential, thought? Do we test sums? I think the book at the back of the room "New Thinking in School Mathematics" might help us a little. On page 17 of that book, there is the statement in the Introductory Address: "There is a need to eliminate dead, useless, outmoded or unimportant parts of mathematics, however hallowed by tradition, in order to develop their understanding and their creative talents." That is a healthy new look at mathematics; mathematics to develop their understanding and their creative talents. A little later in the same book: "It is time we stopped burdening our young pupils with long multiplications and divisions—calculating machines are pretty widespread." There, I think, is a lead into our testing. We do not want to test these rather stupid big sums. We want to test, if we can, the thinking that underlies the mathematics that the child does. We want to cut out, in the main, the routine work, the memory work which still bedevils a good many schools, unfortunately. Mental arithmetic—yes, but it depends. Mental arithmetic is one of those terms which can mean practically anything. In a school I would like to have been to, rather strange things happened in mental arithmetic. Walter Allen, who was a Birmingham schoolmaster turned novelist, wrote in the course of time an autobiography, although he is still only a young man, called "All in a Lifetime", and he has a few words to say on mental arithmetic: "All schoolmasters cultivate some pet foible, some special branch of knowledge to which they attach transcendant importance. Mr. Simmons's was mental arithmetic, but the problems he propounded were based on curious knowledge in which each boy as he came to the school had to be instructed. This was the length of the human intestines. I can see him now at prayers in the big school. No sooner is the last Amen said, than in one prodigious leap, it seems, he comes from behind the lectern to the edge of the dais and pointing a huge, hairy hand at a small boy, he shouts: 'You, boy, the price of your tripes at threepence three-farthings a foot?' " Mental arithmetic means many things to many people.

Marking of tests is a great difficulty. One of the claims made by the devisers of objective tests is that the marking is foolproof. There is a schedule for marking and the answer is right or wrong, true or false. We know in practice that does not always work out. However carefully you devise your marking schedule, there will be errors, there will be answers which have been overlooked. In a subjective type of testing, where the tests have not been, as they call it, "standardized"—and I would rather have called it "sterilized" or "dehydrated" —there are many answers, and the marking schedule is difficult. I had a rather interesting example of this in the 11-plus examination in Birmingham. One question consisted of a graph of the temperature of a classroom throughout a day; it rises and falls, and a certain part of the graph had letters attached, A, B, C. Underneath were certain questions to test whether the child understood the graph. One of the questions was: "What do you think caused the sudden drop in temperature at point B?" The marking schedule was open—any reasonable answer was acceptable. I marked five hundred of these papers and only one gave me any real trouble. A girl, in answering that question, wrote in the space allowed for the answer, "less fidgeting". I had a mental vision of that girl's classroom with the teacher saying, "If you didn't fidget so . . ." I wrote to the examination officer, as there was nothing on his schedule even remotely like that. "What do I do? Is it T for true, is it a tick for right, is it a cross for false, wrong?" To my regret, he sent it back—cross, wrong. I do suggest that the open-end type of question has a most important part to play in our testing, not always these black and white questions. However carefully you devise them, there will be alternative answers.

My next point is the most important of all—the difficult question of speed. This worries me more than any other aspect of testing and teaching. This wretched business of pushing children—"Keep going. Don't stop and look back. Don't reflect. Don't be an intelligent adult. Get on! Get on!" We ought to be training our children, when they have done a little work, to stop, put their pencils down and look at it and think about it. But what happens to them if they do that in these speed tests? The time has gone. They have penalized themselves. We have said from this platform more than once, "go slowly" and yet in most of our testing situations, we are making them just race ahead. It is difficult, and I wish I knew some way out of this. It has the effect in school, when they are doing their practice work, of forcing the teacher to train them for this rat-race sort of stuff. In England, we have been using for many years the Moray House standardized tests. The typical pattern is forty or fifty mechanical sums to be done in 20 minutes. A child who thinks carefully and goes slowly and after each addition sum puts his pen down and goes through it as a subtraction is shockingly penalized, and yet that is the very child with the brains. We have refused to allow any profundity, any depth of thought, to take any place in our testing. That has its effect on the schools. The backwash effect of this speed factor is quite depressing; it is leading to a sort of superficiality in our teaching, to the production of these smart Alecs. Superficially they get on fine, but when they are halted in their tracks and have to think a little bit, they are completely lost.

The excuse given by the devisers of these standardized tests for the speed is "we have got to discriminate, we have got to stretch them into the normal curve of distribution". But have we? Should we test under those conditions? There was some very interesting research which took place in Scotland some years ago into the question of testing, and I would like to read to you some of the findings. I have always held that you should not test a thing until the child knows it, until the child is most likely to get 90-100 per cent right. If you test before that, you are just creating frustration in children. You know what happens with these standardized tests. There are a few at the top with a very high score, a large number in the middle with a moderate score and quite a number at the bottom with 0-10 per cent. Why must we depress and frustrate those children like that? The Scottish Research Council considered this testing business and I shall read you their final statement: "The ideal—and we must accept this is an ideal— the ideal in testing may be stated in the paradox that tests should not be given until the pupils can perform them correctly. To teaching practice in general, it would seem wise to apply the criterion 'teaching or testing?'. If the answer is testing, then we must ask, what is its effect on the pupils?"

Another point about testing which also worries me is that of stereotyping. If you keep to a pattern in your testing—and here again standardized tests tend to—you know the sort of thing there will be; so a lot of teachers cram along those lines, and these tests do lead to cramming, which is an educational evil, an educational curse. The pattern of testing should be regularly and constantly changed. There should be a number of surprise questions every time you test—something that has not been tested before, that has not become stereotyped. Some twenty years ago I went to see a teacher who had applied to become the senior mistress of my school. When I went in to her classroom, they were doing something, and we had a talk and I said: "Well, what are they doing?" "Oh, this lesson is 'intelligence'." They had a lesson in the timetable—intelligence. There they were with those books you can buy in shoals from the booksellers, books of examples of the types of questions you get in intelligence tests. I did not nominate that girl for the job. The problem of stereotyping is difficult—too much change may be harmful, too little change may lead to an undue stress on cramming.

What about the place of words in testing? Here again, we have a very difficult issue. Many years ago, I did some research. I framed a number of arithmetic questions, the same arithmetical situation expressed differently: "What is the difference between 27 and 85?" It can be expressed: "85 take away 27"; it can be expressed: "How many do you add to 27 to make 85?" If you give them the last one, which is exactly the same, a large number of children think that because the word "add" appears, it must be an addition. The wording you use is of the greatest possible importance.

I should like to recount to you a delightful example of the importance of the wording. A girl taking one of these mathematical papers, a girl who had ability, started down the first page and got along fine until she got to about half-way down and then she stuck. Each time the invigilator passed, the girl was still at the same point. She made no progress and at the end of the examination, the invigilator, who knew that the child was a capable child, took her script and looked at it and he found that Mary had reached the question: "Take 8 from 77 as many times as you can." He found written beneath that: "I have taken 8 away from 77 three hundred and sixty-five times and the answer is 69 every time. This is a stupid question." But wasn't it? That girl had done exactly what the rubric asked. She had taken 8 away from 77 again and again and again. Consequently, she got no further in the examination, so that the mark she gained in the paper was no reflection whatsoever on her ability.

So when you make a test, you have got to examine the wording as closely as possible because, with the best will in the world, you can give a very misleading impression.

The next point I wish to discuss is testing in mathematics—what sort of tests may be good and what may be bad—and my tests are all attempting to test the child's power of thinking mathematically. I told you that I had quite a lot to do with the public examination in Birmingham which contained no mechanical sums whatsoever. There was not one single sum in the paper. It was an attempt to assess the children's reasoning powers, their powers of thinking, their acquisition of some of the simpler mathematical concepts. I am talking about that kind of question, and if you are waiting for questions which test sums, then you may leave the room now if you wish. These tests also were used in my school for many years, and let me assure you that you can test a child's mathematical capacity reliably without his doing one single sum. What did we test in these papers? We tested, or tried to test, the fundamentals—notational types, operational types, approximations, estimations, algebraic types, geometry types, translations, comprehensions and so on. I should tell you that, as a result of using this kind of paper for five or six years, the teachers in Birmingham are quite convinced that it is the best kind of test at the present moment. Times will undoubtedly change, but at the present moment, they are very happy for this reason—if they like teaching sums, they can still do them. This does not prevent a teacher from teaching the biggest of sums he likes, but sums will not be tested in the public examination.

I should like to indicate some of the types of questions that we use. I will suggest some that you may find interesting. How do we attempt to test notational ideas? By such questions as: "What is the value of the underlined figure in this number—2,34$\underline{5}$,678?" This is place value, of course. You may think this is easy, but 20,000 Birmingham children are doing these and an astonishing number, even now, cannot give you the place value of a number. Try it with your own pupils and see. Just put a big number on the board, underline one figure, don't go too far along the notational place values, keep within the thousand or ten thousand, and I think you will find much the same. They just have not got one of the fundamentals—the key to calculation. You cannot cope with calculations until you understand notational ideas. If you do not understand these, you should not be calculating.

Some more examples: "What is the difference between the two fours in 404?" One four is four hundred, one four is four ones, difference 396.

"How many times bigger is the first seven than the second seven in the number 7,167?"

"How many tens (or hundreds or tenths or thousandths) can be made from the number 3,456?" Here I would say that if a child divides by 10, then he is stupid. If he understands notation, he will say "three hundred and forty-five tens". He should not have to do a sum.

"Put a decimal point in this string of numbers, 23,456, so that the six is worth six hundredths."

"What is the remainder if I divide 43,362 by 100?" This is obvious. A child, I would submit, ought to say at once: "If you want to know how many hundreds, the answer is 433." If he does a sum, he has not been taught properly.

Take this sort of thing:

363-40

I certainly do not regard that as a sum. If a child does not say: "It is only taking 4 away from the tens column", then he has not been taught properly. The answer is obviously 323; it is only a notational type of question. That holds for many calculations. Some people have looked at these papers I am talking about and said: "Oh, but you have got sums in." I do not call these sums. They are tests of a child's knowledge of the notational system.

Probably the best test of all is to give such a thing as this:

$$6 + 5 = 13$$

How can you make sense of that? Six plus five equals one three. I am not calling it "thirteen" because it is not thirteen. What is the number system, what is the number base that must be operating to make sense of that? I see absolutely no difficulty in a quite young child doing that sort of thing. That, I think, probably is the best type of notational question you can have.

I know some people here are asking: "What's the use of all this stuff? What's the point of working in different bases?" Well, try them and you might possibly appreciate it rather more.

Operational types—these are, of course, in one sense, little sums. But we do not ask them to work out the sum. The important part of a problem is choosing the operation that must be used to get the right answer. When you have done that, it is not worth going any further in most cases. The calculation is usually rather simple; it is the thinking that is important. So I have for many years used questions of that type—a statement, and they are asked to write down just "plus" if it is an addition sum, or whatever you like. I would warn you that if you attempt to give them a problem with two operations, they will find very great difficulty in stating them. This sort of thing: "If 12 pennies weigh 4 oz., what is the weight of 60 pennies?" Two operations there—don't ask them to give you the answer, only: "What two operations are used?" If your children are anything like the ones we have, then they will have considerable difficulty. These are most searching questions. You can give them in the operational type such things as to put in the sign that will be used in a certain statement, this kind of thing—

$$10 \quad \tfrac{1}{2} = 20$$

or

$$24 \div \quad = 6$$

"What must be inserted to make these equations valid?" This can be developed in all sorts of ways.

I have been asked to tell you the other problem I had in Pendlebury's Shilling Arithmetic. On the opposite page to the arrogant English question about the workmen, there was this: "Two monkeys, having stolen a pile of walnuts and filberts from a garden, are on the point of beginning their feast when they see the injured owner of the nuts approaching with a stick. At once, they see that he will take $2\frac{1}{2}$ minutes to reach them. There are twice as many filberts as walnuts and one monkey finishes the latter at the rate of 15 a minute in four-fifths of the time and runs away. The other manages to eat the filberts *just* in time. If the first monkey had stopped to help the other till all were finished, find when they would have got away: (*a*) if they eat filberts at equal rates; (*b*) if the first monkey eats filberts at the same rate as he eats walnuts."

I heard another example of this kind of thing when I was at a course at Gloucester and I had the pleasure of listening to a lecturer give a talk on mathematics in which he recounted this problem which he found in a book published in 1802. "A cardinal can pray a soul out of purgatory in 3 minutes. A bishop will take 7 minutes. A priest will take 12 minutes. How long will it take them praying together to pray five souls out of purgatory?" A rather delightful example of the integration of mathematics into the whole of the school day.

Translation is a mathematical type which has a considerable place in the child's learning. By translation, I mean converting to a verbal statement—again a problem type really—the mathematical statement. I shall give you two that are actually taken from 11-plus papers. The rubric went somewhat along these lines: "Write these sentences in mathematical symbols, *using no words at all:* a gallon of petrol weighs seven-and-a-quarter pounds and a gallon of water weighs ten pounds so the water is two-and-three-quarter pounds heavier than the petrol. Write that mathematically." Here again, of course, the wording is most important. Another one, rather a difficult one for children: "Eighteen less than thirty is the same as five more than seven". That will make them think. I can assure you that only quite a relatively small proportion will effectively translate that statement. I would strongly maintain that that is the essence of the mathematics. If they cannot translate, then they are not beginning to move mathematically There is in these questions almost infinite scope for variation from easy to very difficult indeed.

I have not made any mention yet of the place of money, and this must test the child's ability to use money. I think the only question of importance in the child's work in money in a school is "can he count his money?" I am not interested in sums in money at all, except the simple shopping, domestic types. Try them with questions which are basically counting of money and you will find that they cannot do them. But have we any right to start them doing great big sums in money if they cannot count? It will be an asset to them if they can count their wage packets accurately, won't it? Give them this sort of thing: "Johnny has a money box; he puts coins in his money box and then he empties them out. How much has he got if they were 9 half-crowns, 13 florins, 23 shillings, 14 sixpences?" That is a testing type of question and I think absolutely fundamental. The thing you want to test in money is not the ability to do these:

$$2/11d. \times 6$$

That is the kind of situation that is so utterly false. I do not think it could ever arise in real life except in shopping, and yet we give children these things that are so totally artificial. We come back to what I said the other day—they are only used in school.

Professor Sawyer, who is a professor of mathematics in America, was in a classroom and he put on the board the statement:

$$24 \div 4 = 6$$

He said: "Now children, tell me a story that will show when you would use that mathematical sentence in life. Put it into a social setting." After a considerable time, someone ventured the statement: "We should use it in school, sir!" He said: "Yes, but out of school, when would you ever be likely to use that statement? When would you meet it?" Again a long hush and eventually a child who, I think, wanted to help Professor Sawyer out a little bit, wanted to co-operate; he felt it was not good enough to leave things as they were; he said: "Please sir, mother might send me to the grocer's to buy 24 divided by 4 eggs." Sawyer tells that story in an article he wrote not very long ago, and he apparently replied to the child: "You know, my mother never asked me to do that!"

I think it is useful to bear in mind that most of the maths we teach in school is just left in school. A process in teaching can so easily become teach . . . test . . . forget. Test a group of children who have left school if you do not believe that—and this is under traditional methods. Dr. Land, who is now Vice-principal of Hull University, some years ago gave over 2,000 entrants to training college a relatively easy test in arithmetic and you should have seen the answers, which he published in a pamphlet. One of the questions asked was: "two-thirds multiplied by a half" which is really "what is a half of two?" Nearly a quarter could not do it.

Comprehension, approximations and estimations are three more types of testing situations.

What about simple algebra? Let me tell you another true story about algebra. Some years ago we put into a paper for juniors, questions of this sort:

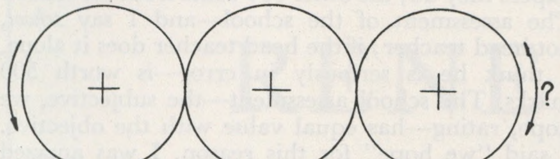
$$3x + 2 = 27$$

What is the value of x to make that true?

These questions had to be submitted finally to the full examination board. They were drawn up by a number of primary school teachers. These papers were made solely by Birmingham teachers, not by administrators, not by inspectors, with due respect to any inspectors who may happen to be here. They were made by Birmingham teachers for Birmingham children. We argued: what do Moray House know about Birmingham? Obviously little. The only test that will suit a Birmingham group of children is a test made by people who know Birmingham children—and that is one of my points for saying that the only real test is a home-made test made by the teachers of the schools, and I would like to see that spreading very much more. Ultimately the paper reached the full examination board which had to approve it. The full examination board consisted mainly of grammar school headmasters. We got to the stage where one of these questions came up and it was debated, discussed. Do we leave this in, do we throw it out? The headmaster of one of our comprehensive schools, a man of great ability, spoke against leaving in the form (3x + 2 = 27). He spoke so ably that when the chairman put the matter to the vote, the vote was overwhelmingly against this question. We primary teachers raised our pencils a little disconsolately to rub it out when the gentleman who had opposed it, said: "If you will put in an asterisk instead of the x, I will agree to it staying in." And it went to Birmingham children in the form (3* + 2 = 27). He was making it more difficult, but it went that way. The next year, when we were framing the questions, we thought: "Right, no x's, let's try capital A, capital B and capital C," and the questions went straight through like that. There was a variety of questions such as:

$$\frac{6 + A}{2} = 12$$

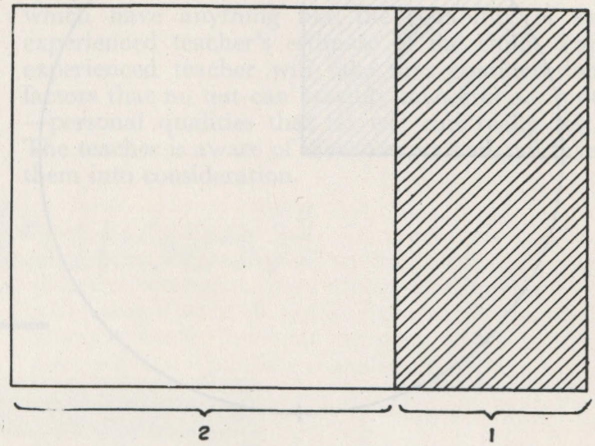
How can we test the geometrical work of a child in the junior school? It is practically impossible, but we have to make an attempt. Any test for children of 11 which does not include some geometry is inadequate. Those of you who read "New Thinking in School Mathematics" will remember that it says in a number of places that in primary school the geometry should be taught on an intuitive basis. How do you test work that is intuitively learned and be fair about it? I think it is almost impossible, but we did include a number of geometrical questions, such as: "If a boy is facing north and he turns round to face east, through what angle has he turned?" One question that we used some years ago—



"Three wheels side by side and touching." (and of course you could extend the number of the wheels if youwished to do so). "Indicate in which direction 'C' turns if 'A' is turned as shown." The sort of thing that the boys on the whole, I think, found rather easier than the girls.

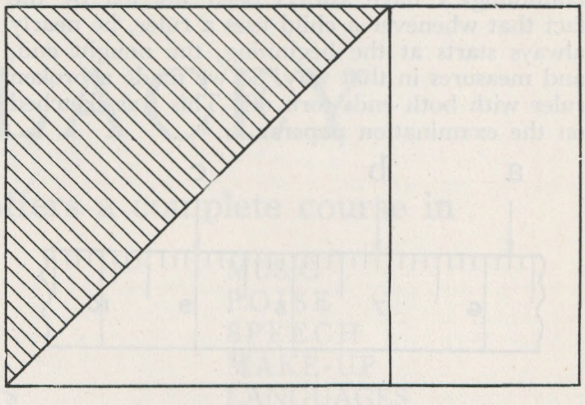
It is so very easy to test this angle stuff; but to me this is of secondary importance in the geometrical work.

There were no sums in fractions at all in these papers. You are no doubt wondering what there was in them. There was enough to stretch the best children in Birmingham. How do we test fractional work? We used largely what we call pictorial fractions. There is no calculation involved at all. This sort of thing—

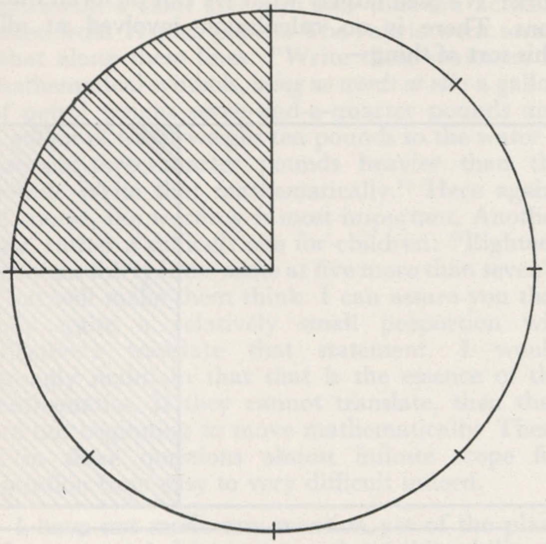


There are clues marked on the diagram and there is a piece shaded. What fraction of the rectangle has been shaded? That is a relatively easy one. The child can mark his paper—the diagram is on the paper.

A rather more difficult one in which once again they must state what fraction of the rectangle has been shaded—



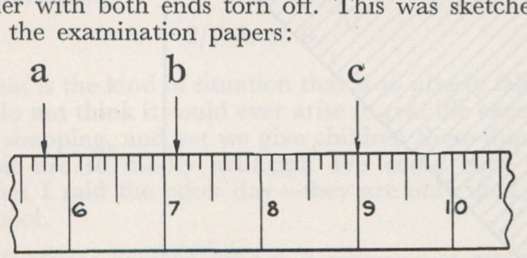
I find these very satisfying. I think you are testing something quite fundamental in fractional work. When you do them, do not stick to one shape; that is a mistake. Give the concept as broad a base as possible. Use triangles, squares, rectangles and circles. Mark the circle to give clues. You have to have clues with this type of question.



This is meant to be marked off into six equal sections. What fraction is shaded? You have a slightly more difficult problem here. You do not go to a point; the child has to think first; it is halfway between them.

They are not easy but they do test the child's acquisition of what I would call the basic fractional notion.

Another type that was found to be extraordinarily resistant to cramming was a measurement sort. However they crammed, it seemed to have little effect on their performance, so they stopped cramming. I have always been worried by the fact that whenever a child uses a ruler, he nearly always starts at the beginning, the nought end, and measures in that way. So we made a broken ruler with both ends torn off. This was sketched on the examination papers:



They saw the markings on the ruler. Certain of the points had letters attached—a, b, c, and questions underneath were of this kind: "How far from a to b?" Try some of these. Children of 11 find them extremely difficult. They have got first of all to say: "What are these divisions? Are they tenths, twelfths or eighths?" Then—"How far from there to there?" You get the most fantastic answers. I would suggest that if a child has not acquired a reasonable mastery of the use of a ruler, we should not give him some of the measurements that we do give him. He has not acquired the use of the basic tool.

An aspect of mathematics that has often puzzled me is time. The only thing of real consequence is "what's the time?" That is our first aim in teaching time, and a lot of youngsters cannot tell the time. Do not give them any sums in time until they can tell the time. You can print clocks and put a time on; not many will get it. You can give questions of this sort and they prove astonishingly difficult to children of about 11: "Johnny left home at 8.46. His walk to school took 8 minutes. After he had reached the playground, he played for 5 minutes, then he went into school. At what time did he go into school?" A child's sense of time is extraordinarily nebulous. I think we should direct some of our attention towards that aspect of our work.

I should like to summarize at this point. In our testing, I would suggest, *we should attempt to test ideas and not facts*. Facts are dead. We should attempt to devise questions which will reduce the effect of cramming. Cramming is a social evil. We can test quite adequately without giving any mechanical sums whatsoever. If we do test, we should also have from the teacher his view of the child's ability—and I would put that as of higher consequence than the score in the test. The Birmingham examination is devised in this way. Let us assume that 1,000 marks are given in the examination. Five hundred of these are gained on the actual papers they do; the other 500 come from the school. The assessment of the school—and I say *school*, not head teacher; if the head teacher does it alone, I think he is seriously in error—is worth 500 marks. The school assessment—the subjective, we hope, rating—has equal value with the objective. I said "we hope" for this reason. I was amazed to find that there were teachers in Birmingham who lacked what I would call professional integrity. For their assessment in school, they gave a range of tests similar to the tests that they had in the examination and based their rating on these tests, which was completely and entirely destroying the whole purpose of the school assessment. My own method was this: I went along to each teacher who had children in the examination and I said: "Put your children in the order of suitability for secondary education of the grammar or technical school type. Do it in any way you like. Put them in order of potential, not order of performance." I also did the same; I put them in what I thought was the order of their potential, their suitability. I had to know them. I did not look at test scores. I find teaching far the most exciting part of being a headmaster and the only way of knowing the children. I did my best to put all of them in order. Then we met and compared ratings, and of course there were marked differences. I may have put Johnny Brown No. 73; his teacher put him very high. We talked about it and we adjusted to get what we thought was a reasonably good list. Eventually, after sweat and toil (and it is much harder than giving a lot of tests and marking them), we were able to send to the education authority the list that we thought was satisfactory. That was the list on which they got 500 marks as a maximum. It seemed to work reasonably well.

Ask the teacher about the child. The thought that the teacher knows better than the test is a most comforting and a most illuminating one. I would like to leave that thought with you. There are no tests, and I do not think ever will be any tests, which have anything like the reliability of an experienced teacher's estimate of the child. The experienced teacher will take into consideration factors that no test can possibly take into account —personal qualities that no test can really test. The teacher is aware of those factors and can take them into consideration.