

Methods of Teaching

Research Evidence vs. Ideology and Self-Interest By Case Vanderwolf

There are many questions in education today that are in urgent need of clear answers. However, the conflicting claims made by the various factions in the debates on education make it difficult for the concerned citizen to distinguish between statements based on sound evidence and those based on personal opinion, ideological conviction, or self-interest. For example, how can we decide if the different methods of teaching children to read (or write or spell or do arithmetic) are all equally effective or whether some methods are better than others? Should a teacher spend most of his or her time teaching the class as a whole or is it better to teach children individually? Do children in small classes learn more than children in large classes? Are there real answers to such questions or are they simply a matter of personal opinion and political conviction?

Consider the argument about the teaching of reading. Is it better to begin by teaching the letters of the alphabet and the sounds they represent (phonics) or is it better to concentrate on reading stories to children while they look at the book that is being read and learn to repeat the story from memory (whole language)? Perhaps it would be better to ask children to learn to recognize entire words, such as CAT or DOG, by their overall appearance but without teaching anything specific about the individual letters that make up those words (whole or sight word method). Still other possibilities can readily be imagined.

A common-sense way of approaching this problem might consist of setting up an experiment in which phonics is tried in one beginning classroom, whole language in a second classroom, and the whole word method in a third classroom. At the end of one or two years, the children in all three classrooms could be given a test consisting, perhaps, of reading aloud from an unfamiliar text or answering questions based on such a text. Suppose that each child is given a percentage score and that the three classrooms achieve average grades of 70%, 50% and 55% respectively. What, if anything, can we conclude? First of all, are the differences among the classrooms real or are they due merely to chance variations? If the tests were given again a week later, would the results be the same?

We can make use of statistics, a branch of mathematics, to help us decide this. Statistical tests can tell us the likelihood that the observed differences are merely random day-to-day variations or an indication of real differences among the three classrooms. These likelihoods are usu-

ally expressed in scientific papers as probabilities (p). Thus, the symbols $p < .05$ mean there are 5 chances in 100 (or less) that the results are merely accidental; $p < .01$ and $p < .001$ mean that the chances of the results being accidental are 1 in 100 (or less) and 1 in 1000 (or less) respectively. If it is found that the results could be due to chance only once in 100 tests, then the differences are usually considered to be real.

However, the fact that the classrooms differ in average achievement does not necessarily mean that the difference is due to the methods of teaching that were used. For example, one teacher may be more successful than another in dealing with young children, regardless of the teaching method used. It is also possible that children in one classroom received more home teaching than the children in the other classrooms or that many children in some classrooms were immigrants. One solution to such problems is to match classrooms at the start of the experiment, attempting to equate such factors as socio-economic status (children from wealthy neighbourhoods tend to perform better in school than those from poorer neighbourhoods), or the years of teaching experience the teacher has had. It is, of course, also essential to be certain that the teachers were actually using the methods they were assigned to use. It is very important to repeat the experiment many times over by assigning a number of classrooms to each of the teaching methods being examined. Experiments based on comparisons between individual classrooms have only a limited significance.

The use of statistical tests to evaluate data from large numbers of subjects is not in any way unique to education research. Very similar methods are used in field trials of drug treatments, or other procedures in medical practice. For example, in evaluating a drug that may be useful in combating high blood pressure, thousands of patients taking the drug (in various doses) may be compared with patients who are given no treatment or a control treatment (perhaps a pill containing only a little sugar or starch).

Patients may improve when given a control treatment merely because they believe they are receiving an effective treatment (placebo effect). Similarly, in education research, both students and teachers may perform better if they are involved in the testing of a new method, as opposed to a traditional one, regardless of the relative merits of the two methods (Hawthorne effect).

Although the problems involved in education research are complex, they are not at all insurmountable. There is an enormous body of work involving large-scale comparisons of different methods of teaching children how to read, write and spell. The results are very clear. Teaching that relies on systematic phonics results in better reading accuracy and better comprehension than teaching that relies on whole language or

whole word methods (Chall, 1983); Engelmann et al., 1988; Stahl & Miller, 1989). Whole language methods produce inferior results even when they have the benefit of novelty (Usher & Evans, 1978). The most recent research evidence indicates that young children appear to benefit from having stories read to them, but actual instruction in reading should begin with the letters of the alphabet and the sounds they represent, followed by extensive practice in reading simple stories (Adams, 1990). The research evidence that is currently available also suggests answers to many other questions in education policy. For example, studies in which self-esteem was assessed by psychological tests indicate that self-esteem is enhanced by competence. Children who have learned to read, write, compute, etc. think better of themselves than children who cannot do these things. Attempts to improve school performance by artificial attempts to inflate self-esteem are generally ineffective (Engelmann et al., 1988).

Similarly, there is much evidence that teaching methods whereby the teacher sets the pace, sets the standards, and teaches the class as a whole tend to achieve better results than teaching methods that emphasize student-directed learning and individual "conferencing" with students (Gutierrez & Slavin, 1992; Schoen, 1978; O'Neill, 1988). If the teacher is preoccupied with one student, the rest of the class is unsupervised and the time devoted to the lesson is reduced.

Although it is often claimed that large classes inevitably result in inferior education, a careful examination of the facts available suggests a different conclusion. A two-year study in Toronto, in which the effects of class sizes of 16, 23, 30 and 37 students were compared in a well-designed experiment showed no effect of class size on written composition, vocabulary, reading and mathematical problem-solving in grades four and five (Shapson et al, 1980). This result is consistent with much other work showing that class size has very little relation to student achievement (Coleman et al, 1988; Shapson, 1973; Thorndike, 1973) except, perhaps in the case of very small classes (Glass & Smith, 1979).

Educators today feel themselves under attack by parent groups who question not only the methods and standards of instruction in the public schools but also the adequacy of the training received by teachers and the necessity for the large non-teaching bureaucracy that has developed in education (Nikiforuk, 1993). It is perhaps only natural that people under attack will attempt to defend themselves as best they can. When questioned, educators will often assert that child-centred, activity-based education policies are based on strong scientific evidence; that children's literacy skills are greater now than in past generations; that whole language methods have worked exceptionally well in New Zealand, England or the Orient; and that increases in the pupil-to-teacher ra-

tio will destroy high-quality education. Although none of these assertions has any basis in fact, it is always best to remain open to reasonable persuasion. Ask to see the evidence. If no evidence is forthcoming (invariably the case in my experience with the education bureaucracy), it is justifiable to ask persistently, loudly and frequently why the present policies are being followed and what is being done to change them.

(Dr. Vanderwolf is professor of Neuroscience at the University of Western Ontario.)