Lake Wobegon Revisited:
On Diversity and Education

ANDREW BIELMILLER

I present the theses that current educational practices underestimate the magnitude of diversity in educational achievement and that disregarding diversity has the effect of widening the range of differences. I suggest that diversity in the rate of growth of capacity to process increasingly complex tasks partly underlies achievement differences. I propose a transactional model of the interaction of capacity differences and curriculum that results in failure to achieve academic skill mastery by unnecessarily large numbers of children. I propose that reducing concern with "grade-level achievement" and increasing concern with academic mastery would have long-term educational benefits for individuals and society.

Educational Researcher, Vol. 22, No. 9, pp. 7-12.

In Garrison Keillor's mythical Lake Wobegon, "all the children are above average." When the idea is put this way, we immediately see the impossibility of such a state of affairs. Yet similar objectives have been accepted as part of the rhetoric of educational criticism. In the influential book, A Place Called School, Goodlad wrote, "But let us assume . . . that virtually all students can be brought along a full grade per year by techniques involving peer assistance, diagnostic testing, corrective feedback, and the like" (1984, p. 166). The simple truth is that no one has yet demonstrated methods that begin to approach bringing "virtually all students" along a full grade level a year (i.e., bringing virtually all children to the current average level of performance). While there is doubtless room for considerable improvement in school learning, all available evidence suggests that there will continue to be a wide range of performance. By sixth grade, at least 30% of children will continue to lag a grade level or more below the "average" child and still further below "above-average" children.

Recent writers (e.g., Good & Brophy, 1987; Goodlad, 1984; Oakes, 1992; Radwansky, 1988) and authorities (e.g., National Governors' Association, 1990) have suggested approaches to dealing with diversity in schools that do not involve teaching children differently as a function of their ability. The main thesis of this essay is that educational approaches to diversity that ignore skill differences may perpetuate these differences at current levels. While it is certainly true that at present, many less successful children get short shrift in school (less time, poorer instruction, etc.) as Barr & Dreeben (1983), Hiebert (1983), Oakes (1992), and others have described, I will argue that commonly proposed remedies—abolishing ability grouping and emphasizing cooperative learning—may further disadvantage those already disadvantaged by society. I will suggest that while some observed diversity is clearly the product of environmental disadvantage, much diversity is also observed in environmentally advantaged populations. I will further suggest that educational practices that ignore the reality of developmental diversity in achievement may actually amplify the impact of naturally occurring educational diversity. Finally, I will propose that educational programs that acknowledge diversity provide the best hope for minimizing its impact on skill acquisition.

My goal is to achieve an educational system that provides a more "level playing field" than many children now experience. I should be clear that my focus is on skill development in elementary schools. By and large, I agree with critics of secondary education (e.g., Goodlad, 1984; Oakes, 1992) who find programs excessively "tracked" and frequently irrelevant in content. I might add that other changes are probably needed in what Oakes calls the "technology of instruction." What I am proposing here are changes in the "technology of instruction" in elementary school that might result in a relatively more equal secondary school population.

The Magnitude of Diversity in Educational Skill Acquisition

The actual magnitude of observed educational skill diversity in our society is great—invoking a range of more than four grades by grade 4. Consider the magnitude of differences in items passed on a standard test of reading and mathematics skill. Figure 1 shows percentages of mathematics items passed on the Wide Range Achievement Test (WRAT; Jastak & Wilkinson, 1984). The same items were administered to all 5- to 12-year-old children who participated in norming this test. These items reflect computations of increasing difficulty from single digit addition and subtraction to complex operations with fractions. The test is administered individually. Multiple choice responses are not used. Figure 1 was derived from the "Level 1" norms for children between 5 and 12 years of age. "One hundred percent" was arbitrarily taken as the number of items passed by 12-year-old children in the 50th percentile. At that age, 90th percentile children pass approximately 35% more items than the 50th percentile children. Tenth percentile children pass less than 60% of items passed by 50th percentile children. (Note that this test is not timed.)

Another way of looking at the same data is to note that the level of performance on mathematics attained by the 10th percentile children near 12 years of age was attained by the 90th percentile children when they were less than 8 years old.

Andrew Biemiller is professor, Institute of Child Study, University of Toronto, 45 Welles Road, Toronto ON MSR 2X2, Canada. He specializes in education, children's self-direction, and reading.
old. The data given here refers to raw scores. We can legitimately suggest that children are demonstrating similar levels of achievement at similar raw score levels.

The data for the Reading (word recognition) scale of the WRAT are similar. Thus, in both word identification and computation, the level of skill reached by 50th percentile 10-year-olds on the WRAT is reached by 90th percentile 8-year-olds and 10th percentile 12-year-olds.

Examination of growth on individual standardized test items yields similar conclusions. Examination of published item norms for the Canadian Test of Basic Skills (King, 1982) indicates that on given items, increases in accuracy on reading comprehension items averages about 20% across three grades while growth on math items averages a little over 30% across three grades. To illustrate, correct responses on an item testing two digit subtraction with regrouping increase from 40% in grade 3 to 75% in grade 5. Note that this means that “independent mastery” of this skill cannot be considered a “pre-requisite skill” for all children in any of these grades, although the level of mastery is clearly improving. These findings show that there is really a very wide range of performance among children in schools.

SES and Ethnic/Visible Minority Diversity

There has been much concern about socioeconomic status (SES) and ethnic/visible minority differences in educational achievement (see, e.g., Good & Brophy, 1987; Goodlad, 1984; McGill-Franzen & Allington, 1991; Oakes, 1992). These writers suggest that many SES and ethnic/visible minority differences are attributable to school practices, and they stress the need to provide compensatory experiences for disadvantaged children. I agree.

Too often, when children are initially grouped by ability in order to avoid asking the “less advanced” children to do the impossible, the “less advanced” are then taught too slowly. For example, evidence from “Reading Recovery” research (Pinnell, 1989), analyses of pace of instruction (Barr & Dreeben, 1983), text difficulty (Chall & Conrad, 1991), and the effects of assisted reading practice (Shany, 1992) all support the view that many children have the potential to master skills faster than they do now in school. These findings imply that instructional practices with more slowly progressing children need to change.

Furthermore, it is undoubtedly the case that lower SES and minority children are more likely to be included in “less advanced” ability groups. While Haller (1985) found evidence that group assignments were not “biased” in terms of assessed ability, minority children were at greater risk of having lower assessed abilities when groups were being formed. The question is whether these children will benefit from instruction geared to those who have acquired more skills. Chall & Conrad (1991) recommended that when reading with teacher assistance available, all readers, including below- and above-average readers, use texts that are about one grade level above their own tested reading levels. Given the ranges of diversity, this would involve creating different groups using different texts. Hiebert, Colt, Catto, and Gury (1992) reported that “at-risk” (Chapter 1) children withdrawn from class and given “restructured” reading instruction starting at their ability levels reached the mean level of reading performance attained by children with somewhat higher reading readiness scores who received ungrouped “whole language” classroom instruction. (We do not know how the latter group would have fared given the “restructured” educational program provided to the “at-risk” children.) Hiebert et al.’s “restructured” group showed substantially more diversity (as indexed by standard deviations) on all outcome measures than a group with initially comparable readiness scores. Similarly, Madden, Slavin, Karweit, Dolan and Wasik (1993) report that 3 years of cross-age ability-grouped reading instruction (with focus on phonics and basal readers), combined with individual tutoring, family support, and other features, resulted in average gains of 1.0 grade equivalents in decoding (Woodcock Word Attack), 0.7 grade equivalents in letter and word identification (Woodcock Letter and Word Identification), and 0.5 grade equivalents for oral reading (Durrell Analysis of Reading Difficulty). There was an increase in diversity compared with an initially comparable control group for the decoding measure, but not for word identification or oral reading.

Findings in these experimental studies are consistent with reviews showing that within-class ability-grouped math instruction and cross-age ability-grouped reading instruction ("Joplin Plan") had beneficial effects on the achievement of both less skilled and more skilled elementary school students (Gutiérrez & Slavin, 1992; Slavin, 1987). (No systematic evidence is available regarding within-class ability grouping for reading.) In short, it appears that high-quality instruction geared to student ability levels in specific skills results in greater gains for at-risk elementary school students than does heterogeneously grouped skill instruction. However, there is some suggestion that such instruction may also lead to greater diversity of outcomes, as some children gain

---

FIGURE 1. Range of Achievement: Math

---

8 EDUCATIONAL RESEARCHER
more than others from enriched curricula. Clearly, use of high-quality instruction is desirable, but equally clearly, it will not reduce diversity and may increase it.

Children from advantaged backgrounds (and from countries where there is less cultural and income diversity than in North America) also show wide diversity in school performance. For example, Stevenson et al. (1990) report the same magnitude of within-grade variance between children in their Japanese, Taiwanese, and American samples. Thus, while there were significant mean achievement differences between children from the three different countries, there was no evidence that instructional methods leading to

The actual magnitude of observed educational skill diversity in our society is great—involving a range of more than four grades by grade 4.

greater mean achievement reduced individual variation. Much of my own work has involved SES-advantaged children in a laboratory school in Toronto. Some of our children, from “advantaged” and loving homes, have the misfortune to be born “young” (i.e., in November or December), to mature a little more slowly, or to lack talent in academic areas. The achievement range of the laboratory school population covers approximately two thirds of the total range indicated on standard achievement tests. In this advantaged population, progress in reading comprehension is strongly related to underlying developmental variables such as naming speed for letters (Biemiller, 1977/1978). Similarly, children in this school who are performing well above grade level in mathematics have IQs that are 20 points higher than children performing below grade level.

Even if effective compensatory approaches were to eliminate SES/minority differences, there would still be a substantial range of performance among children of the same age. When we confront teachers and school systems with expectations that all children should be performing at or above age-normed “grade levels,” we ensure “failure” for a significant proportion of children or teachers. Furthermore, if we insist that children from all schools in a district should show similar distributions of abilities, we ignore the impact of poverty and prejudice on children’s development. When we set unrealistic expectations, we create an atmosphere in which teachers, principals, and board officials may respond with practices that obscure the truth (Haladyna, Nolan, & Haas, 1991; McGill-Franzen & Allington, 1993; Smith, 1991).

Growth in Capacity to Deal With Increasingly Complex Tasks

What is the nature of the underlying capacity that increases with maturity and permits increasing levels of educational achievement? I suggest the capacity to deal with increasingly complex tasks. Modern theories of cognitive development (Case, 1985; Fischer & Pipp, 1984; Flavell, 1972; Siegler, 1981; Pascual-Leone, 1970) all use a “cognitive capacity” construct in the form of increasing working memory. This is an analogy to the random access memory (RAM) in a computer. Robbie Case (1985) provides many illustrations. For example, a task can require simple coordination of two components—placing one block on top of a second, while placing a third block next to the second. At around 18 months, a child can perform one component or the other. But the child cannot do both together. By 24 months the two components can be combined. Growth in the capacity to “think about more things at one time” can also be illustrated with a “working memory” task. On the backwards digit scale of the WAIS (Wechsler, 1958), adults experience problems reversing more than four or five digits. On the children’s version of the test, WISC (Wechsler, 1974), the majority of 6-year-olds experience problems when reversing more than two digits.

Sequences of Complexity

In short, children’s capacity to deal with “complexity” increases with age. But what is “complexity”? In practical terms, it refers to the number of “things” the child must “think about” or “be aware of” when doing a task, and the “abstractness” of those “things.” For example, compare carrying out simple addition versus addition with regrouping. More conscious steps are involved in regrouping.
The basic point is that just as children’s mental growth can be described in terms of increasing complexity, so the tasks we give children can be ordered in terms of increasing complexity.

As Case (1978, 1985) points out, the sequence of curriculum broadly reflects increasing complexity demands. This is most clearly demonstrable in mathematics. However, in reading, vocabulary and narrative complexity also increase (Chall, Jacobs, & Baldwin, 1990). There is a similar increase in the complexity of narrations children can be expected to generate orally or in print (McKeough, 1992). To the extent that the later elementary school skills curriculum requires a greater level of capacity to deal with complexity, those who have developed that capacity at the point when they encounter the curriculum will be in a better position to profit from it.

A Transactional View of the Development of Expertise

How do the child’s capacity for dealing with complexity and the complexity demands of a curriculum in a specific domain interact to produce an independently competent child versus one who lacks confidence and competence? In a recent article, Biemiller and Meichenbaum (1992b) propose that cognitively advantaged children (children who for reasons of age, experience, and rate of maturation can deal with relatively higher levels of complexity than their peers) spend much of their time in schools working on tasks that are “well-matched” or easy for them. The regular school curriculum gives them tasks that leave them with “surplus cognitive capacity.” Operationally, this means they could learn somewhat more complex tasks than those taught. We hypothesize that this “surplus capacity” is used to “verbalize” tasks. Such verbalization reflects a higher level of verbal self-regulation by these children.

A considerable body of modern educational theory (e.g., Bereiter & Scardamalia, 1986; Berk, 1992; Brown & Palincsar, 1989; Diaz, 1992; Peterson, 1988; Vygotsky, 1978) suggests that higher levels of verbal self-regulation permit more effective problem solving. My colleague, Donald Meichen-
baum, and I have observed that during independent classroom work, teacher-nominated “high self-directed” children produce twice as many task-regulative statements as “low self-directed” children (Meichenbaum & Biemiller, 1992). Berk (1992) summarizes similar findings in other studies. In a more recent analysis of these data, we found that the “self-directed” children helped others with their tasks about once every 15 minutes during “independent work” periods. This contrasts with once per hour for “low self-directed” children (Biemiller & Meichenbaum, 1992a).

We suggest that through these task-regulative experiences, cognitively advantaged children learn (a) that they are capable of mastering what they have to learn and (b) how to verbalize tasks to themselves and to others whom they assist. They become able to “take charge” of their own learning in ways that other children are not observed to do. Consequently, these children acquire both basic performance competence and a form of “elementary expertise” for curriculum tasks.

Cognitively disadvantaged children quite simply have insufficient cognitive capacity to handle the complexity demands of some or all school tasks as assigned. They constantly require assistance and learn more about how to get assistance than how to do tasks independently. Proposals for heterogeneous grouping and cooperative education that result in some children usually receiving assistance in academic matters, and others usually receiving such assistance, appear likely to reinforce dependency. (It is important to emphasize that a child may be considerably more advanced in one domain than another. Bloom, 1976, and others document this point.)

Our observations indicate that a much wider range of children are capable of demonstrating “elementary expertise” on academic skills. In both classrooms (Meichenbaum & Biemiller, 1992) and an experimental cross-age tutoring context (Biemiller, Shany, Inglis, & Meichenbaum, 1993), “low self-directed children” markedly increased levels of task-directive speech when given academic tasks of reduced complexity.

Consequences of Demands That All Children Carry Out the Same Work at the Same Time

Giving the same tasks to all children in a grade, as Goodlad (1984) found was routinely done in mathematics and some reading instruction, means that individuals at different levels of cognitive development reach very different levels of mastery of tasks (if they “master” them at all). Chall and Conrad (1991) also document this outcome in their study of text difficulty and comprehension. As we have seen in previous sections, there are several insidious problems of defining “success” as a specific level of performance at a specific age. These include (a) problems for those who simply “can’t,” and (b) problems for those who barely “can.”

The Effects of Asking Children to Do What They Can’t

These are well illustrated in John Holt’s classic How Children Fail (1959). Motivationally, we create a situation analogous to demanding that all play basketball like 6’6” basketball pros. One can compare this approach to academic learning with the old issue of school sports for enjoyment versus “competitive teams.” Most of us don’t “make the team.” The majority anticipate failure when success is defined as what the 10% or 20% “best” do (see, e.g., Atkinson, 1965; Dweck, 1986).

The Effects of Asking Children to Do What They Barely Can

A second problem of defining success as equal performance at the same age is its effect on those who just barely “can.” Not only do those who can’t perform a specific task at a specific grade “fail,” but those who just barely can—who can complete the work sheet, solve the problem, and so forth—are also deprived of effective learning. When performance requires children to use their full resources, leaving no “surplus capacity” to talk to themselves and others about

---

Our observations indicate that a much wider range of children are capable of demonstrating “elementary expertise” on academic skills.

---

the task, knowledge about the performance may not be verbalized and may consequently be unavailable when really needed. This suggestion is admittedly more speculative. Most of the evidence for it comes from the “cognitive instruction” literature, examining the quality of private speech or task-directive speech while working on school tasks (see, e.g., Berk, 1986, 1992; Diaz, Neal, & Amaya-Williams, 1990; Meichenbaum & Biemiller, 1992; Peterson, 1988). In a recently completed study (Biemiller et al., 1993), classroom observations indicate that the amount and quality of task-directive speech of peer-nominated “middle self-directed” children is virtually identical to that observed in peer-nominated “low self-directed” children, and much lower than that of “high self-directed” children. In short, those who can just “keep up with” curriculum in terms of passing performance tests may also be at risk of not learning skills well enough to use them when needed in new situations.

Implications

It appears that the conditions experienced by advanced children foster their expertise. It is a plausible hypothesis that creating these conditions for children who now fail or barely succeed could advance their expertise as well. However, doing so may involve variations in rates of progress, as suggested by Carroll (1989).

Creating and Operating Expertise-Supporting Programs

I recognize that the type of teaching implied by the transactional perspective involves a fundamentally more complex educational program than is now common. I do not propose to design an entire curriculum in this article, nor do I believe there is a single program design that flows from this perspective. However, I can suggest some general characteristics of programs that are designed to meet the learning needs of a wider range of children. As in ungraded programs (Goodlad & Anderson, 1963), children would progress in various skills at varying rates. Achieving mastery of skills would be a higher priority than would conforming to a fixed curriculum schedule. These programs
will probably involve either use of split grades/family grouping, or team teaching organized across two or three classrooms with different grades. From the child’s viewpoint, I do not see instruction in new skills occupying a large portion of time—perhaps an average of 30 or 40 minutes a day, usually in homogeneous groupings. (Teachers might spend as much as 2 or 3 hours in introducing new skills to small groups.) Settings for practice or consolidation of skills in different domains would probably occupy approximately half of the child’s time. Assisted practice settings (assisted by the teacher or more skilled peers) might alternate with independent practice settings on different days for different children. Expertise settings, in which children (a) assist younger and less skilled children and (b) apply well-consolidated academic skills to “meaningful” activities, are important for all children, not just the most able. It may be important to ensure that cognitively advanced children do not foreclose expertise experiences of others. An hour a day in expertise settings would be a large increase over what is now experienced by the majority of children. In essence, the intent of these suggestions is to create the conditions now experienced by “advanced children” for a much larger proportion of children in a classroom.4

Summary

The range of performance across elementary school children is immense. For various academic domains, the level reached by children near the bottom of any class (10th percentile) in grade 6 is reached by children near the top of the class (90th percentile) some 4 to 5 years earlier.

If we teach using methods that treat everyone as though they were at the same level—when they are not—children at the low end of the class will have one of two experiences—failing or just barely making it. Both of these will have negative consequences for their sense of competence in the school domain they are studying. By contrast, if we change the way in which we treat elementary school children, and move to a system more analogous to that used by the Red Cross for its swimming badges (i.e., ascending levels through which all children move at their own pace, until they receive a particular color badge, which signals complete mastery), all children would become proficient at some level in all subjects. Perhaps, more important, all students would have the experience of starting out at the bottom of a level, and then slowly rising to the top, where they would have surplus capacity and be able to engage in a variety of metacognitive activities that are not possible when one is failing or just making it.

What is proposed here is similar to Goodlad and Anderson’s (1963) proposals for “non-graded schools.” As Gutiérrez and Slavin (1992) have recently summarized, comparisons of such programs, when implemented on a group, rather than an individualized basis, generally showed substantial advantages for children from all ability levels. Slavin (1987) reported similar results in an earlier review of ability grouped instruction. (It should be noted that there is some overlap in the studies reported in these reviews.) What is proposed here differs from these earlier approaches primarily in placing greater emphasis on providing “less advanced” students with the kinds of “mastery experiences” now enjoyed by more advanced students.

This sort of system would be better for the children who are not well served by the present one and just as good for those are currently well served. The objective would be to have more children leave elementary school with a high level of competence in literacy and numeracy skills. Thus, it would represent a better overall utilization of our human resources.

Notes

I wish to thank several individuals who have advised me on the writing of this article, including Jeanne Chall, Ellen Regan, Don Meichenbaum, and the anonymous reviewers of the two drafts, especially reviewers 2 and 4 on the second draft. Of course, these individuals are not responsible for the content, but have helped greatly in clarifying the main points of the article and avoiding irrelevant points. I also wish to thank the children and staff of the Laboratory School of the Institute of Child Study, who have contributed to my understanding of individual and developmental differences for 20 years. Finally, I must express an intellectual debt to Willard Olson, whom I have never met, for first introducing me to the problem of developmental diversity in education.

Much of the information presented in this section is not “new.” However, although diversity is a pervasive fact in nonstreamed classrooms, texts directed at teacher education candidates do not always reflect this reality. For example, in the excellent and widely used Looking in Classrooms (Good & Brophy, 1987), four chapters (4, 9, 10, and 11) address issues intimately related to diversity. However, no quantitative information on ability ranges is provided. Contradictory advice against ability grouping (when possible) and for maintaining a close “match” between ability and learning tasks is provided on a number of occasions. The same can be said of widely used texts in educational psychology (e.g., Gage & Berly, 1988). In my personal experience working with teachers over 20 years, I find that most are surprised by the magnitude of diversity described in this article.

There are other literature reviews which do not support “ability grouping” in any form (e.g., Berg, 1966; Esposito, 1973; Findlay & Bryan, 1971). Slavin (1987) argues that there is evidence that group instruction has beneficial effects at all ability levels for elementary school students who are “regrouped” for a small portion of a day by the specific ability tested, taught at levels commensurate with their abilities (e.g., not simply the “third grade math curriculum”), and assessed with instruments that reflect the range of skills instructed. Slavin also stresses that “students should remain in heterogeneous classes at most times and be regrouped by ability only in subjects (e.g., reading, mathematics) in which reducing heterogeneity is particularly important. Students’ primary identification should be with a heterogeneous class” (1987, p. 328). I completely agree with this final point.

It is argued by some that learning is not necessarily sequential. This proposition is true if we are deciding whether to teach Shakespeare before Herringway—or whether either must be taught. The proposition is clearly false if we propose to deal with large numbers in mathematics before dealing with small numbers. In early literacy instruction, there are legitimate debates about the criteria to be used in selecting vocabulary for reading, spelling, and general knowledge. However, once a vocabulary is established, it is highly desirable to let students build on that vocabulary, rather than confronting them with new materials containing too many words that are unfamiliar in print or meaning (Barr, Sadow, & Blachowicz, 1989; Chall & Conrad, 1991). In short, “complexity sequences” reflect both “logical” sequences of skills in which later concepts require grasp of prerequisites, and “content” sequences in which we continue to use knowledge we have acquired and add to it. These ideas were described in somewhat more detail in Bimmiller and Meichenbaum (1992a).

References


DECEMBER 1993
11


Received August 17, 1992
Final revision received May 30, 1993
Accepted June 21, 1993