Educational Evaluation and Policy Analysis
Fall 1995, Vol. 17, No. 3, pp. 323–336

Opportunity to Learn and Conceptions of Educational Equality

Gretchen Guiton
University of Southern California
Jeannie Oakes
University of California, Los Angeles

Focusing on the equity aspect of proposals for making opportunity-to-learn standards integral to an accountability system, this article discusses conceptual issues surrounding determination of equal educational opportunity and explores ways that these issues manifest themselves in empirical formulations of opportunity to learn (OTL). Using two databases, OTL measures are developed according to three alternative conceptions of equality—the Libertarian, Liberal, and Democratic Liberal conceptions—and the influence of these conceptions on the information provided is compared. This examination shows the intimate relation between values on equality and measures of equality and brings these issues to the fore for discussion by educators and policymakers.

For 30 years, opportunity to learn has been a potent concept for researchers conducting international comparative studies, allowing them to conceptualize and measure varying aspects of the relationship between teaching and learning—for example, the amount of time required for students to complete a task or the degree of overlap between content taught and content tested (Anderson, 1990). In the past decade, opportunity to learn has become a controversial policy issue stemming from the notion that measures of schooling processes, including teaching and learning, can play a critical role in national reform efforts aimed at more effective and equitable schools (Traiman, 1993).

In this article, we discuss knotty conceptual issues in developing and using opportunity-to-learn (OTL) standards to inform policy questions of equal educational opportunity. We begin by examining recent trends in educational reform that provide the context in which OTL has gained center stage. We then critically examine current proposals for making OTL standards integral to an accountability system. Focusing on the equity aspect of these proposals, we discuss conceptual issues surrounding determination of equal educational opportunity and explore ways that these issues manifest themselves in empirical formulations of OTL. In this examination, we attempt to show the intimate relation between values on equality and measures of equality and to bring these issues to the fore for discussion by educators and policymakers.

The Reform Context

More than a decade ago, reform reports (e.g., *A Nation at Risk*, 1983) declared a nationwide decline in educational achievement of crisis proportions. These reports mobilized public support for reforms that increased inputs (e.g., lengthening the school day) and raised standards (e.g., increased graduation requirements) in an effort to encourage schools and students to redouble their efforts (Goodlad & Keating, 1990; Smith & O'Day, 1990). These top-down reforms, and the legislative and financial support they commanded, quickly led to a call for greater public scrutiny of education and a renewed interest in the development of educational indicators to monitor reforms (see

e.g., Burstein, Oakes, & Guiton, 1992; National Commission on Educational Excellence, 1983; Selden, 1988). For example, standardized testing of both students and teachers increased dramatically. These early reform efforts focused mainly on minimum academic competencies and 'more of the same' as a way to improve education.

By the late 1980s, the limitations of these efforts became apparent, and policymakers and educators began instead to address fundamental expectations about student learning and instructional practices, and, with them, to restructure relations among administrators, teachers, parents, and students—moving the initiative to those closest to instruction (Elmore, 1990). Knowledge of change processes, the limited impact of top-down reforms, and the strong local-control tradition in education reinforced the desire to provide maximum local flexibility, authority, and responsibility. However, some, questioning the generalizability and long-term success of such highly localized reforms, advocated supportive, cohesive policy structures to achieve the radical reconceptualization of schooling considered necessary to meeting the educational needs of a changing society (Smith & O'Day, 1990).

Accountability, especially conceived of relative to outcome measures of student and school performance, became a critical, central component of these proposed changes (Murphy, 1993). Concerned with the need for high levels of performance, national education goals were adopted and standards promulgated to provide a "yardstick against which students, parents, teachers, and others could measure individual and system progress toward the Goals" (NCEST, 1992). Explicit national standards for student achievement and a voluntary system of assessments that emphasize higher order thinking and complex problem-solving skills were seen as the primary mechanisms for achieving systemic change (Resnick, 1992).

The use of student outcome indicators presumes agreement on the nature of achievement—that some limited outcomes represent appropriate goals for all students and schools—and that these outcomes can be validly and reliably measured (Bryk & Hermanson, 1993; Oakes, 1989). Moreover, the proposed accountability uses of outcome indicators—to administer sanctions and incentives based on individual student and school performance—presume an educational system that provides adequate human and material resources to ensure a fair opportunity for students to achieve these outcomes and that schools' efforts predominantly account for student performance. Although some agreement about broad educational goals and the benefits of high standards and assessments to measure progress toward them exists,[1] assumptions about equitable learning opportunities and schools' control over learning outcomes lack credibility, and proposals to link student or school outcomes to sanctions or incentives have met with strong resistance.

Critics of these proposals point to the considerable research evidence documenting the inequitable distribution of educational resources and access to knowledge. For instance, finance equalization efforts have failed to close the gap among and within school districts (Elmore & Fuhrman, 1993; Kozol, 1991). The least qualified teachers teach minority and poor students (California Commission on the Teaching Profession, 1985; Darling-Hammond, 1990, 1995). Likewise, students in low-income, high minority schools have less access to computers, equipment, and laboratories (Oakes, 1990; U.S. Department of Commerce, 1989). Perhaps most important, given the strong relation between course taking and achievement, are the inequities that exist in access to curriculum. Curriculum differentiation—the practice of making different knowledge available to different groups of students—sometimes signifies responsiveness to students' needs, but raises questions about unequal opportunities as well (Oakes, Gamoran, & Page, 1992). Studies of secondary schools report that low-income and high minority schools offer fewer advanced courses and low-income and minority students are disproportionately assigned to vocational and remedial tracks (Oakes, 1985, 1990; Oakes & Guiton, 1995). Each of these areas—quality teachers, materials and equipment, and tracking—demonstrate significant distributional inequities, relate to student achievement, and are

not fully in the control of the school. According to critics of outcome indicators, these resources and processes represent those warranting consideration before holding individual students and schools accountable.

Opportunity-to-Learn Standards

As a consequence of these concerns about the uneven distribution of educational resources and access to knowledge, recent proposals advocate including some form of OTL standards in the proposed system of content and outcome standards, to mitigate against the unfairness of high-stakes testing of students who attend inadequate schools (e.g., NCEST, 1992; Goals 2000, 1994). Educators and policymakers troubled over the unforeseen and oftentimes negative impact of high-stakes testing support these proposals. They caution that accountability systems that use student achievement data in isolation would penalize schools serving less advantaged students and function as a strong disincentive for educators to teach in those schools (Darling-Hammond, 1994).

These concerns prompted much heated debate as Congress considered inclusion of OTL standards in Goals 2000: Educate America Act. Proponents won a victory when Congress included a provision requiring states that receive Goals 2000 funding to develop OTL standards. Opposition to the measure was powerful enough, however, to weaken the requirement considerably, in that, although states are required to develop OTL standards, any review of them is voluntary.

In the act, OTL standards are defined as

> the criteria for, and the basis of, assessing the sufficiency of quality of the resources practices, and conditions necessary at each level of the education system (schools, local educational agencies, and states) to provide all students with an opportunity to learn the material in voluntary national content standards (Conference Report, 1994).

This definition highlights the important link between OTL and outcomes and, thereby, addresses advocates' concerns that OTL standards are needed to assure fairness in a high-stakes, outcome-based accountability system. Students and schools, advocates argue, should not be held accountable for performance in the absence of adequate resources and learning conditions. Opponents counter that OTL standards will stifle local creativity, further burden educators, and, as a consequence, work against the interests and outcomes of students (Traiman, 1993).

Characterized in this way, the debate over OTL disguises the underlying complexity and deep-rooted beliefs that shape both sides of this issue. The division over OTL reflects fundamental differences in views of the appropriate role of federal and state policy (and their historical limitations) to assure equity; disputes about the utility of standards and assessment to influence practice; basic disagreements about the nature of the educational system and how it works; and competing conceptions of what is required to obtain educational equality (see e.g., Bryk & Hermanson, 1993; Darling-Hammond, 1994; Elmore & Fuhrman, 1993; Guiton, 1992; Howe, 1994; Oakes, 1995; Porter, 1995; Strike, 1991).

The Relation Between Accountability and OTL Standards

Much recent literature on OTL standards addresses these issues of definition, utility, and relation to other policy conceptions such as equity and effectiveness. These accounts incorporate historical and theoretical evidence in thoughtful discussions about appropriate considerations for proposing or using OTL standards. Three examples illustrate the considerable variation among these discussions.

O'Day and Smith (1993) proposed a conception of accountability that included recommendations for OTL standards. Defining "a quality education as the opportunity to learn well the content of the frameworks," school standards derive from a definition of educational quality and the principle that all students should have access to it. Thus, linking quality and equality, these standards "provide operational specifications for assessing whether students in the school have the opportunity to learn the content and skills set out in the frameworks" (p. 274). Standards include resource, practice, and performance criteria closely (and narrowly) aligned to content frameworks. O'Day and

Smith specify four ways in which standards might be used to ensure equal opportunity to learn while improving curriculum and instruction: ongoing self-generated improvement efforts; periodic system-generated monitoring or review mechanisms; the use of rewards and sanctions for students and schools; and informing judicial and legislative decisions.

In a similar vein, Porter (1995) distinguishes between accountability and the school improvement purposes of OTL standards. Porter suggests that both student and school accountability be accomplished by focusing on student performance in core academic areas through student assessments carefully aligned to content standards. He takes this approach to accountability because he is unsure that OTL standards would motivate failing schools and because he believes that waiting for schools to meet OTL standards prior to implementing student testing would only result in low standards or an indefinite wait (1995).

Despite his skepticism about the value of OTL standards for accountability, Porter sees them benefiting school improvement efforts. Such standards, he argues, could support a system of school process indicators that provides descriptive information about instruction. This system, in turn, would provide empirically supported visions of good practice, communicate models, and inspire improvement. These two uses, accountability and improvement, differ considerably. Accountability uses force action in prescriptive ways, whereas improvement uses offer guidance and provide information. Consequently, policymakers must be clear about the intended uses to guide criteria for certifying voluntary standards.

Darling-Hammond (1994) proffers a different notion of accountability based on the premise that schools have a "duty of care"—that is, they are obligated to "treat students well and responsibly, to provide them equal access to educational opportunity, to adhere to professional standards of practice, and to use the best available knowledge in developing appropriate strategies for teaching each child" (p. 192). Accountability requires that practices and policies be continually evaluated and revised. In this way, accountability efforts support educators in their task of discovering and adopting good practices and support policymakers in their task of creating structures and resources that promote the use and availability of such practices (p. 196). Based on this conception of accountability, Darling-Hammond proposes three types of standards: (a) standards that address state and district responsibilities for providing equitable access to learning opportunities and, at the school level, (b) practice standards (goals and general approaches for school practices), and (c) standards for accountable functioning (how schools manage, self-evaluate, and problem-solve).

Whereas O'Day and Smith recognize and support a range of uses for OTL standards, Porter sees these uses conflicting and advocates that policymakers select between accountability and improvement uses. Darling-Hammond, in contrast, proposes an affirmative responsibility for monitoring the quality of educational services as well as their distribution. Accountability requires processes for guiding practice, so she distinguishes the level in the system where OTL standards should be monitored with the majority at the local level.

Given these differing perspectives on accountability, it's not surprising that even those policymakers and educators who see value in the concept of opportunity to learn have difficulty reaching agreement about what valid and useful OTL indicators might be. Although much of the rhetoric on OTL standards addresses their use in assuring equity—however defined—OTL indicators also provide information on the quality of schooling experiences. By definition, OTL indicators describe the resources, school conditions, curriculum, and teaching that students experience, even though their ostensible purpose is to assess differences among students in these "opportunities." Policymakers and researchers are confused about what this descriptive information about the nature of schooling processes might mean in an accountability system, as well as about the role it should play and how salient it should be. These multiple uses of OTL indicators pervade the accountability uses described previously.

Even when we focus only on the equity uses of OTL indicators, considerable ambiguity remains about how OTL indicators should be conceptualized and used. Should primary emphasis be placed on using OTL indicators to portray the inequalities in students' school experiences? Or should OTL indicators be used primarily to condition our explanations of inequalities in students' outcomes? How should equality be defined? What standard should be used in determining whether or not equity obtains, and what role should national, state, and local educators have in making that determination? What empirical evidence is needed to establish whether or not equal educational opportunity standards have been met? Should the primary concern be to assess the extent to which all students have the same opportunities—whether the playing field is even? Should attention also be paid to the nature of these opportunities—what the playing field is like? Should judgments be made about whether available opportunities are adequate to accomplish various goals of the education system? Is OTL important primarily as an intervening variable, that is, to help us better explain why particular outcomes obtain? Who makes these determinations? What are the consequences of these choices for policy and practice?

OTL: Empirical Analyses and Conceptions of Equity

In the remainder of this article, we explore these questions further using empirical evidence. Recognizing that the development of indicators is itself a value-laden activity, we illustrate the influence of three alternative conceptions of equality on the development of OTL measures, and on the kinds of information that such OTL indicators provide (Burstein et al., 1992; de Neufville, 1975). Three coherent, commonly-ascribed-to ideological positions—the classical liberal or Libertarian position, the equal opportunity or Liberal position, and the Democratic Liberal or egalitarian position—dominate modern substantive theories of the appropriate distribution of social goods (Ericson, 1990; Howe, 1995). These positions on equality lead to consideration of different elements in the educational system and set forth distinct criteria for determining that equality has been achieved. For these three positions, we ascertain whether differences in position result in consistent differences in determinations of what educational components warrant equitable distribution and how they should be allocated. Thus, we look at the way different ideologies manifest themselves in OTL indicator data and use these differences to address fundamental issues in the development and use of OTL standards.

Two separate databases inform these investigations. One national database includes information on the U.S. eighth-grade sample collected by the International Association for the Evaluation of Educational Achievement during the 1981–1982 school year. The public school sample [Second International Mathematics Study (SIMS), Population A] contains information on 93 districts, 126 schools, and 234 classrooms, and more than 5,000 students; however, these data should not be interpreted as being nationally representative because of difficulty in obtaining district participation (Travers, 1993).[2] The second database includes 12 years of longitudinal student data for 3 public school systems involved in recent desegregation litigation. Analyses are reported separately for each district.[3]

We limit our analyses to equity in track placement, the major mechanism by which students are provided differentiated curricular and instructional experiences. And given that critical track placement decisions are made about students in middle school, we focus on that level. Substantively we ask, from each ideological position, how equitably distributed are track placements among middle school students? We use analyses to reflect on ways that an indicator system may be designed to inform policy and practice.

The Libertarian Position. The Libertarian, adopting a formal equality perspective, focuses on the processes by which one acquires or transfers educational goods. Thus, it holds that inequality in the distribution of these goods is warranted when differences are made in a procedurally fair process on the basis of relevant characteristics of the recipients—that is, merit. In education, choice,

tenacity, and ability often are considered educationally relevant characteristics justifying unequal distribution (Green, 1980, p. 49). School resources and opportunities, in this view, could be unevenly distributed among students whose educational qualifications differ, with those with higher qualifications deserving greater resources or "better" opportunities. At the same time, this position supports only minimal involvement of the state in the distributional process. Therefore, the position advocates a minimalist approach to judging merit and allocating resources that prohibits intrusion of the state into such "private" matters as the race, gender, or social class of recipients, or how educationally relevant characteristics intertwine with these ascriptive characteristics in the distributional process.

Now, a Libertarian might wish to investigate the equity of student participation in mathematics classes using data from the SIMS study and from the three desegregating school districts. Doing so, the Libertarian, permitting differentiation of opportunities based on relevant characteristics, would ask, "Is there a positive relationship between ability level and placement in higher track classrooms?"

If the Libertarian were satisfied with understanding whether, on average, students in the higher track classes initially scored better on measures of mathematics achievement than did students in the lower track classes, (that is, to examine the central tendencies in these data), he or she would be pleased with the distribution of curricular opportunity, represented by track level. For example, an analysis of variance (ANOVA) of the SIMS data, using the mean score of each class on a mathematics pretest as an estimate of ability, demonstrates that higher level classes (i.e., higher tracks) contain, on average, those students with higher ability levels.[4] Similarly, the data used in the Rockford, San Jose, and Wilmington desegregation cases show significant differences in the average scores of students in different math tracks, with the lowest mean score in the lowest track and the highest mean in the high track. From these results, the Libertarian would conclude that class assignment followed a meritocratic process and thus was equitable.

However, a Libertarian might reach different conclusions were he or she holding the educational system to a standard of placing *every* student in math classes according to merit. Using this standard, then, the Libertarian would examine the ranges and variance within classes and the extent to which classes at different levels had overlapping score ranges. Overlaps in the math abilities of students in classes at different track levels would lead the Libertarian to conclude that, at least to some extent, class assignment, by deviating from a strict meritocratic process, lacked procedural fairness and, therefore, was unjust.[5]

With this higher standard, both databases would raise concerns for the Libertarian. In the SIMS data, not only was there substantial variation in students' math abilities within classes, but the overlaps among classes at different levels also was considerable. For each class type, the third quartile class mean was at least as high as the first quartile class mean of the more advanced class. The median score for "enriched" classes was higher than that of the first quartile in the higher level algebra classes. Thus, 50% of the class means for "enriched" classes were above the first quartile algebra class means. These analyses show that characteristics other than the immediately relevant measure of merit appear to be influencing the placement process. Equally meritorious eighth graders seem not to have equal access to math classes at various levels.

Because the SIMS data use classes from a large number of schools, it is not possible to ascertain whether these inequalities are a function of a flawed process within particular schools or whether they result from the vagaries of different schools' placement practices. It may be that these data simply show what Garet and DeLany term the *frog pond effect*, with different schools applying locally determined standards of merit (Garet & DeLany, 1988). If it turned out that schools with many highly able students were setting higher standards for admitting students to high-level classes than were schools with few high-scoring students, a Libertarian might conclude that, although there may be some systemwide inequalities at work, no particular school was employing an unfair practice.

Data from the three court cases permit a finer grained analysis, because the analyst can consider data about each district and school separately. Doing so in San Jose, for example, he or she would find that sixth graders placed in a low-track mathematics course demonstrate abilities ranging from rock-bottom Normal Curve Equivalent (NCE) achievement scores on the mathematics battery of Comprehensive Test of Basic Skills of 1 to extraordinarily high scores of 86. Even more striking, achievement scores of sixth graders in standard-track math classes span the entire range, from NCE scores of 1 to 99. And, although sixth graders in accelerated courses have a somewhat more restricted ability range, they too score all the way from 52 to 99 NCEs. Thus, this school system appears to place a very large number of students according to criteria other than merit (as measured by standardized achievement tests),[6] particularly those students with scores between 52 and 86. Students with these NCE values could be found in any of the three math class levels.

Using such standards and measures to assess students' opportunities to learn, the Libertarian would conclude that inequalities mar the educational process. However, the minimalist approach to state intervention would prevent the Libertarian from going farther to ascertain whether or not this inequitable distribution of opportunity is a function of particular irrelevant characteristics, such as students' race, gender, or social class.

The Liberal Position. The Liberal, taking an equal opportunity position, shares the Libertarian's concern for a merit-based process and accepts an uneven distribution of resources and opportunities on that basis. But, unlike the Libertarian, the Liberal takes a more direct stance that looks to ensure that an uneven distribution is based on a fair competition—that is, it does not result from irrelevant characteristics such as race, social class, and gender. The Liberal also supports a more activist role for the state that includes compensating for disadvantage to even the playing field and help insure a fair competition. Consequently, both the resources provided and the procedure by which those resources are distributed warrant scrutiny in analyses of equality. In addition to employing standards of merit for any distributional inequality, the Liberal standard for judging equality also requires that neither the quality of the resources nor the processes by which they are provided relate to irrelevant student characteristics such as gender, ethnicity, or family socioeconomic status (SES).

Thus, before accepting a conclusion about whether track-level assignments reflect equality or inequality, the Liberal would ask whether assignment to class depends *solely* on relevant criteria—such as measured ability—or whether it depends to some extent on irrelevant criteria such as gender, ethnicity, or family SES. Using our two data sets, the Liberal would address this equity question by examining the relationships among measured ability, background characteristics, and class placement directly.

With regard to the SIMS data, for example, a Liberal would probably want to examine our three analyses of covariance using class level (track) as the dependent variable; gender, ethnicity, or SES as the independent variable; and class mean pretest score as the covariate.[7] Although SES did not account for differences in track placement once pretest scores were considered, both class ethnic and gender composition were related to track,[8] suggesting that the irrelevant characteristics, ethnic and gender composition, influenced placement over and above the influence of ability. Because these analyses found greater access to high-level courses for females and minorities, and the Liberal position supports the distribution of additional resources to disadvantaged groups, the Liberal would conclude that equality exists in track placement and, in this case, the unequal distribution of access to courses would be considered justified as a kind of compensation.

If the Liberal were satisfied with information that on *average*, nationwide, a student with "advantaged" background characteristics had no greater likelihood than comparably meritorious "disadvantaged" students of being provided access to higher level math classes, then these analyses would be satisfactory. However, here, too, the Liberal with another standard—one concerned about

whether *individual* students are being judged only by relevant criteria—might be cautious about this conclusion based on aggregate data.[9]

The "equality" result in the SIMS analysis might, in fact, be an artifact of differences in "cut" scores that determine merit in different schools and systems. One could not rest assured that females and minorities were actually being favored in comparison with equally achieving males and White students. It could be, for example, that most of the minorities included in the database attended predominantly minority schools where both the overall level of math achievement and qualifying scores for entry into higher track classes were quite low. Even in such schools, however, Whites may have been advantaged in their access to higher level math courses over comparably scoring minorities.[10]

Analyses of data from the three court cases would prove more illuminating to the Liberal holding this individual standard. As a group, African American and Latino students scored lower on achievement tests than Whites and Asians in all three of the school systems. However, in each of the systems, African American and Latino students were much less likely than White or Asian students *with the same test scores* to be placed in high-level courses. For example, in San Jose, Latino eighth graders with "average" scores in mathematics were three times *less* likely than Whites with the same scores to be placed in an "Accelerated" math course. Among ninth graders, the results were similar. Those Latinos scoring between 40–49, 50–59, and 60–69 NCEs were less than half as likely as their White and Asian counterparts to be placed in accelerated tracks. The discrimination was even more striking among the highest scoring students. Whereas only 56% of Latinos scoring between 90–99 NCEs were placed in accelerated classes, 93% of Whites and 97% of Asians gained admission to these classes. Rockford and Wilmington data exhibit similar patterns. In all three systems course placement practices skewed enrollments in favor of Whites and Asians over and above that which could be explained by merit, as measured by tests of achievement.[11] Such outcomes would cause those holding a Liberal position considerable concern.

The Liberal would want to go further than simply scrutinizing the procedural fairness of the placement process when considering the equality of such placements. He or she would also be interested in ascertaining whether, once placed in classes at different levels, students had equal access to resources and opportunities likely to promote or constrain future opportunity. These resources and opportunities might be defined as access to qualified teachers, to various topics and skills, and to various instructional strategies. For example, the SIMS data permit investigation of the relationship between teacher resources, content coverage, and instructional goals. Here we discuss these differences for classes of different racial composition.[12] We develop indices of teacher experience (average number of years teaching and teaching eighth grade), education (number of math and math-pedagogy courses), and assignment (proportion of teaching assignment devoted to math classes). Classes predominantly comprised of Whites and Asians had higher levels on all three indices of teacher quality than mixed or predominantly minority classes, although these differences were not statistically significant when ANOVAs were conducted.[13] Correlations between new content coverage in five topic areas (fractions, ratios, measurement, geometry, and algebra) suggest that significantly more fraction and ratio subtopics are "introduced" as minority composition of the class increases. In addition, when examining teachers' content emphases, we find that they hold significantly fewer goals for predominantly minority classes in fractions and ratios than they do for predominantly White and Asian classes. In combination, these results suggest that teachers expect mixed and predominantly minority classes to have less previous exposure to these topics, warranting treating more topics as new material, yet simultaneously holding lower expectations for students mastering the content—a result that would concern the Liberal as it suggests that some classes, identifiable on the basis of an irrelevant characteristic, receive more low-level math content, yet teachers of these classes hold lower expectations for student achievement and, themselves, are likely to be less qualified.

The Liberal would also be concerned that class placements constrain future opportunities, a pattern suggested by Rosenbaum (1986), who argues that access to high-status curricula is maintained only by a series of student "wins" (demonstrations of ability, effort, and achievement), and any loss removes the student from consideration for these opportunities. Moreover, even students who win in the low-status curriculum are prevented from moving up, if they have missed experiences considered prerequisite to a higher curricula—a circumstance suggested by the content coverage analyses reported previously (Hallinan, 1987; Oakes, 1985).

The Democratic Liberal Position. The Democratic Liberal, taking an egalitarian position, goes beyond requiring the absence of barriers related to irrelevant characteristics to take a more substantive approach, requiring a redistribution to obtain some threshold level of performance. Note that this position goes beyond compensation to requiring redistribution until the performance level is met: This position incorporates outcomes in the equality analysis and therefore, appears most like what is being called for in national discussions of OTL standards, but its strong redistributionist requirements probably exceed many policymakers' intent (see Howe, 1994; Oakes, 1995).

In contrast to the Libertarian and Liberal positions that base judgments about the equity of resource and opportunity allocation in relation to procedural fairness and relevant or irrelevant characteristics of the recipients, the Democratic position takes a substantive view of equality. It first establishes, either normatively or empirically, a threshold level of benefits that applies to all students and then places an affirmative responsibility on the state to assure that all students attain this minimal level. Hence, it relegates abilities to the same arbitrary category as social differences and places a far greater emphasis on whether the distribution of resources and opportunities situates recipients equally to achieve the established standard.

Following from this interest in equal access to outcomes, the Democratic position first asks of our data: To what extent do various classes successfully help their students reach the threshold with the criterion being that all students should attain this level? Because participation in algebra 1 by the ninth grade has long been considered a key indicator of later access to and participation in 4-year colleges and universities, we treated readiness for algebra by the ninth grade as a threshold standard in the SIMS data.[14] We operationalized a standard for readiness for algebra coursework empirically, deriving it from the pretest performance of students in eighth-grade algebra classes. We reasoned that students who were placed in algebra in eighth grade had met their school system's determination of algebra readiness. We determined the proportion of students in each remedial, general, and enriched class who met the threshold[15]—valued numerically as the first quartile of mean pretest scores of the algebra classes—on the posttest, and, thereby, could be considered algebra-ready by the end of the eighth grade. This indicator, then, could help the Democrat to judge the extent that different class placements are equal in their impact for various groups of students.

Analyses of the SIMS data show differences in the mean proportion of students meeting the threshold by class type (14% in remedial classes, 34% in general classes, and 70% in enriched classes). Analyses by class composition indicate no differences by gender composition, but show significant differences between high SES and low, low-moderate, and moderate SES classes and between predominantly White and Asian classes and predominantly minority and mixed classes. These results would indicate to the Democrat that race and socioeconomic characteristics represent "needs" warranting additional resources.

Analyses of the data about students in the cities engaged in desegregation litigation might take a slightly different tack. These longitudinal data reveal the rates of growth for students with comparable past achievement as a function of being assigned to math classes at different levels. This type of value-added analysis presumes minimally that all students benefit equally from their class placement—that is, that relative disadvantages not become worse (Howe, 1994). For

example, in San Jose, students who were placed in lower level courses consistently demonstrated lesser gains in achievement over time than their peers placed in high-level courses. For example, among the students with preplacement math achievement between 50 and 59 NCEs, those who were placed in a low-track course began with a mean of 54.4 NCEs, but lost an average of 2.2 NCEs after 1 year, and had lost a total of 1.9 NCEs after 3 years. Students who scored between 50 and 59 NCEs and were placed in a standard-track course, by contrast, began with a mean of 54.6 NCEs, gained 0.1 NCEs after 1 year, and had gained 3.5 NCEs after 3 years. The largest gains were experienced by students who were placed in an accelerated course. They began with a mean of 55.4 NCEs, gained 6.5 NCEs after 1 year, and had gained a total of 9.6 NCEs after 3 years. These results were consistent across achievement levels: Whether students began with relatively high or relatively low achievement, those who were placed in lower level courses showed lesser gains over time than similarly situated students (i.e., those with similar pretest scores) who were placed in higher level courses. Again, we found similar patterns in Rockford and Wilmington. Like the SIMS analyses showing that classes at different levels provided students with dramatically different prospects for achieving a threshold standard, these analyses would warrant the judgment by those holding the Democratic position that class assignments were inequitable in the three districts.

Implications for Standards and Indicators

What do these illustrative analyses imply for the development and use of OTL indicators? Clearly, the three positions lead to consideration of different elements in the educational system. The Libertarian focuses exclusively on resources, the Liberal expands consideration to schooling and classroom processes, whereas the Democratic position considers inputs and processes primarily as they influence the outcome standard—achievement of the threshold. Even when all three positions consider the same element (here placement in class type) application of the standard associated with each position can lead to very different determinations. Equality of what is a critical question (Sen, 1992). Because equality is not one cohesive belief that applies in all spheres, our conception of what is to be distributed and the value we ascribe to it—in this case education—relates closely to the function we see it fulfilling in our society, the amount of state activism we will tolerate, and the principle of equality we embrace. These differences result in varied criteria used to determine whether equality applies (Verba & Orren, 1985).

Whether OTL indicators are used to hold schools accountable, diagnose problems, or monitor the status of the system, their use requires agreement about the degree of equality indicated. These analyses suggest that fundamental tensions exist in conceptions of equality that prohibit such agreement and that these differences have practical as well as theoretical importance. Unless these tensions are recognized and explicitly addressed in indicator studies, we doubt that much progress will be made toward achieving equity.

How might such OTL indicators contribute to goals of equal educational opportunity? The use of multiple measures, representing different equality perspectives, can clarify conflicting conceptions of equality. When determinations of equality in educational opportunity result, agreement of indicators, representing consensus across positions, can garner support for decisions. When indicators jointly confirm inequities, they are likely to be most powerful as the sense of injustice provides a cogent starting point for social assessment (Shklar, 1990, in Sen, 1992). However, OTL indicators frequently will provide conflicting or ambiguous results; in these cases, they can shed light on differences exposing them for scrutiny and discussion. Only through such open efforts to address norms that permit the egregiously uneven distributions of resources and opportunities that exist today are we likely to make any progress.[16] After all, public disclosure lies at the heart of accountability in a democratic system.

Such discussion, focused on what is happening, why, and whether existing practices

are accomplishing what is intended, is likely to be most effective at the school or district level when conducted as part of an ongoing collective dialogue (Darling-Hammond, 1994). If such reflective conversations are to occur, we need mechanisms for collecting and reporting such information back to schools and districts. Proposals for school-based feedback mechanisms suggest that schools will require support, encouragement, and external pressure to reflect on data and their implications for changing practice (Clune, 1993; Darling-Hammond, 1994; Oakes, Sirotnik, & Hare, 1986). Many recognize the role of OTL indicators in arousing public interest, opening up new ways of considering problems, and deepening public discourse (Bryk & Hermanson, 1993; Darling-Hammond, 1994; de Neufville, 1975; Oakes, 1989, 1995; O'Day & Smith, 1993). By keeping our conceptions of equity in the forefront of our measurement efforts, we enhance the possibility of fairly representing pluralistic interests, of recognizing and considering competing values in our decisions, and of providing tools to aid an ongoing, dynamic effort to consider what we want from our schools and to accomplish those ends.

Notes

This article was prepared for *Educational Evaluation and Policy Analysis*, issue dedicated to the memory of Leigh Burstein. Authors, listed alphabetically, contributed equally to the article.

[1]Although highly visible reform efforts seem to agree in their focus on higher order thinking for all students, we would caution that this is only one of the many outcomes addressed by schools.

[2]To control for level of analyses problems, all analyses using the SIMS sample were conducted on class-level data. The class level was chosen because resources and opportunities are distributed to classes rather to individual students.

[3]The first system, Rockford Public Schools, in Rockford, Illinois (previously under an interim court order) was the target of a federal liability suit brought by a community group, The People Who Care. Among other complaints, the group charged the school system with within-school segregation through ability grouping and discrimination against the district's African American and Latino students, nearly 30% of the total enrollment. The other two systems, San Jose Unified School District, in San Jose, California, and New Castle County Schools in Wilmington, Delaware, approached the court, hoping to be released from long-standing federal desegregation orders. The plaintiffs in both of these cases argued, among other things, that the districts had used their ability grouping systems to create within-school segregation and, thereby, circumvented the intent of their court orders.

The analyses used here are reported more completely in Jeannie Oakes's 1993 report to the court in *The People Who Care v. Rockford Board of Education School District No. 205*, in her July 1993 deposition in conjunction with *Jose Vasquez et al. v. San Jose Unified School District et al.*, and in a January 1995 report to the court in *Coalition v. New Castle County Schools*.

The Rockford system was found liable by the court in 1993. The San Jose system reached a settlement prior to the formal hearing date in 1993. Both systems are under court order to alter their grouping practices. The judge has not yet ruled in the New Castle County case.

[4]In the SIMS data, class type refers to four distinct course offerings (i.e., remedial, general, enriched, and algebra) and corresponding curricular differentiation. Post-hoc Scheffe tests indicate significant differences ($p < .01$) between mean ability levels for all class types.

[5]Note that these analyses take student achievement on a standardized test as the sole basis for determining merit. Although the authors, and many readers, may not agree with this use of test scores, it is analogous to proposed uses of outcome measures in indicator systems.

[6]Most schools attempt to track students according to prior achievement usually using a combination of achievement scores, grades, and teacher recommendations (see e.g., Oakes & Guiton, 1995: Oakes, Gamoran, & Page, 1992).

[7]To control for level of analyses problems, all analyses are conducted at the class level. Student demographic variables were aggregated to the class level and represent classes with high concentrations on that variable. For gender, classes that were 65% or higher in males and 65% higher in females were compared to classes having equal numbers of males and females. For SES, classes were assigned to one of four levels based on a combination of parent education and occupation. Students' reported ethnicity was used to identify predominantly minority, mixed, and predominantly White or Asian classes. Three separate analyses of covariance (ANCOVAs) were conducted, because policy interest exists for each irrelevant characteristic separately and because data were analyzed at the class level.

[8]The class mean for predominantly female classes is higher than that for predominantly male

classes, and both are significantly higher than the class mean in gender-balanced classes. Mean class type for predominantly White and Asian classes is lower than that for either the mixed or predominantly minority classes. This counterintuitve finding may result from lower achievement criteria for high-track placement in predominantly minority schools.

[9]There is considerable ambiguity in support for group or individual rights—whether group membership serves as the basis for recognizing individual deprivation or both the basis for identification and the target of remediation as in affirmative action (see Guiton, 1992, pp. 20–22).

[10]See Oakes & Guiton (1995) for data that demonstrate these placement patterns.

[11]Because each of these school systems had undertaken efforts and had reasonable success in desegregating their school buildings (a voluntary effort in Rockford, and court-ordered efforts in San Jose and New Castle County Schools), the distribution of Whites and minorities was fairly even across schools. This condition combined with desegregated analyses that permitted us to examine the data for individual schools underlie our confidence that the systemwide patterns described here are not an artifact of aggregating data across very different schools.

[12]Numerous other process variables were investigated using these data: Drawing on educational research, items in the SIMS database were combined to assess input and process constructs considered important in the literature (e.g., multiple indicators of teacher, organizational, and instructional resources, and student resources and demographic characteristics, class composition, content coverage, instructional practices, and classroom environment). The examples given are merely illustrative. See Guiton (1992) for a complete description of the data and methods of analyses.

[13]Using another measure of relationship (the averaged implicit weights, which are based on the ratio of the average amount of the good held by members of the special group—here minority classes—to the average amount held by members not in that group—predominantly White and Asian classes), all indicators, except mixed/White and Asian for education, are less than 1, indicating inequality. These differences with the results of the ANOVAs demonstrate the ambiguity that arises from using different measures to determine equality (Guiton, 1992). It is well-documented in the economics and educational finance literature that the degree of equality is often ambiguous depending on the particular measures used (see e.g., Berne & Stiefel, 1984; Sen, 1973, 1992).

[14]This standard derives from the widespread interest in short-term academic outcomes and from the emphasis on higher levels of mathematics literacy for all (e.g., Murnane & Raizen, 1988). Additionally, algebra at the ninth grade serves a major gate-keeping function for college course work (e.g., Oakes, 1990).

[15]This value was slightly lower than the first standard deviation below the mean of algebra classes on the pretest.

[16]See Kozol (1991) for a description of the significant inequalities that exist in schools. Also see Elmore & Fuhrman (1993) for a description of one and a half centuries of failed state efforts to achieve equality in educational funding.

References

Anderson, L. W. (1990). *Opportunity to learn and the National Assessment of Educational Progress: An analysis with recommendations*. Unpublished paper.

Berne, R., & Stiefel, L. (1984). *The measurement of equity in school finance*. Baltimore: Johns Hopkins University.

Bryk, A. S., & Hermanson, K. L. (1993). Educational indicator systems: Observations on the structure, interpretation, and use. In L. Darling-Hammond (Ed.), *Review of Research in Education: Vol. 19* (pp. 451–484). Washington, DC: AERA.

Burstein, L., Oakes, J., & Guiton, G. (1992). Educational indicators. In M. C. Alkin (Ed.), *Encyclopedia of educational research: Vol. 2* (pp. 409–418). New York: Macmillan.

California Commission on the Teaching Profession. (1985). *Who will teach our children?* Sacramento, CA: Author.

Conference Report on H.R. 1804. Goals 2000, Educate America Act. (1994). *Congressional Record, 140*(32), H1625–H1684.

Clune, W. H. (1993). The shift from equity to adequacy in school finance. *The World & I, 8*(9), 389–405.

Darling-Hammond, L. (1990). Institutional policy into practice: "The power of bottom over the top." *Educational Evaluation and Policy Analysis, 12*, 233–241.

Darling-Hammond, L. (1994). Standards of practice for learner-centered schools. In Berne, R., & Picus, L. O. (Eds.), *Outcome equity in education* (pp. 191–223). Thousand Oaks, CA: Corwin.

de Neufville, J. I. (1975). *Social indicators and public policy*. Amsterdam: Elsvier Scientific.

Ericson, D. P. (1990). *Social justice, evaluation, and the educational system*. Unpublished paper. Honolulu: University of Hawaii.

Elmore, R. (1990). Introduction: On changing the structure of public schools. In R. F. Elmore

(Ed.), *Restructuring schools* (pp. 1–28). San Francisco: Jossey-Bass.

Elmore, R., & Fuhrman, S. H. (1993). Opportunity to learn and the state role in education. In S. L. Traiman (Ed.), *The debate on opportunity-to-learn standards: Supporting works* (pp. 73–102). Washington, DC: National Governors' Association.

Garet, M. S., & DeLany, B. (1988). Students, courses, and stratification. *Sociology of Education, 61*, 61–77.

Goals 2000: Educate America Act of 1994, Pub. L. No. 103-227, § 1-3, 108 Stat. 125 (1994).

Goodlad, J. I., & Keating, P. (Eds.). (1990). *Access to knowledge*. New York: The College Board.

Green, T. F., with Ericson, D. P., & Seidman, R. H. (1980). *Predicting the behavior of the educational system*. Syracuse, NY: Syracuse University Press.

Guiton, G. (1992). *Developing indicators of educational equality: Conceptual and methodological dilemmas*. Unpublished doctoral dissertation. Los Angeles: UCLA.

Hallinan, M. T. (1987). *The social organization of schools: New conceptualizations of the learning process*. New Haven, CT: Yale.

Howe, K. R. (1994). Standards, assessment, and conceptions of equality. *Educational Researcher, 23*(8), 27–33.

Kozol, J. (1991). *Savage inequalities: Children in America's schools*. New York: Crown.

Murnane, R. J., & Raizen, S. A. (1988). *Improving indicators of the quality of science and mathematics education in grades K–12*. Washington, DC: National Academy Press.

Murphy, J. (1993). Restructuring: In search of a moment. In J. Murphy & P. Hallinger (Eds.), *Restructuring schooling* (pp. 1–31). New York: Teachers' College.

National Commission on Excellence in Education. (1983). *A nation at risk*. Washington, DC: U.S. Government Printing Office.

National Council on Educational Standards and Testing. (1992). *Raising standards for American education*. Washington, DC: Author.

Oakes, J. (1985). *Keeping track: How schools structure inequality*. New Haven, CT: Yale.

Oakes, J. (1986). *Educational indicators: A guide for policymakers*. Santa Monica, CA: The RAND Corporation.

Oakes, J. (1989). What education indicators? The case for assessing the school context. *Educational Evaluation and Policy Analysis, 11*, 181–199.

Oakes, J. (1990). *Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science*. Santa Monica: The RAND Corporation.

Oakes, J. (1995). Opportunity to learn: Can standards-based reform be equity-based reform? In the 75th Anniversary Commemorative Volume of the National Council of Teachers of Mathematics.

Oakes, J., Gamoran, A., & Page, R. (1992). Curriculum differentiation: Opportunities, outcomes, and meanings. In P. Jackson (Ed.), *Handbook of research on curriculum* (pp. 570–608). New York: Macmillan.

Oakes, J., & Guiton, G. (1995). Matchmaking: The dynamics of high school tracking decisions. *American Educational Research Journal, 32*, 151–181.

Oakes, J., Sirotnik, J. A., & Hare, S. E. (1986). Collaborative inquiry: A congenial paradigm in a cantankerous world. *Teachers College Record, 87*, 545–562.

O'Day, J., & Smith, M. S. (1993). Systemic reform and educational opportunity. In S. H. Fuhrman (Ed.), *Designing coherent education policy*. San Francisco: Jossey-Bass.

Porter, A. C. (1995). The uses and misuses of opportunity-to-learn standards. *Educational Researcher, 24*(1), 21–27.

Resnick, L. B. (1992). Standards, assessment, and educational quality. *Stanford Law and Policy Review, Winter*, 1992–1993, 53–59.

Rosenbaum, J. E. (1986). Institutional career structures and the social construction of ability. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education*. New York: Greenwood.

Selden, R. (1988). Missing data: A progress report form the states. *Phi Delta Kappan, 69*(7), 509–513.

Sen, A. K. (1973). *On economic inequality*. New York: Norton.

Sen, A. K. (1992). *Inequality reexamined*. New York: Russell Sage Foundation.

Shklar, J. (1990). *The faces of injustice*. New Haven, CT: Yale.

Smith, M. S., & O'Day, J. (1990). Systemic school reform. *Politics of Education Yearbook*, 233–267.

Strike, K. (1991). The moral role of schooling in a liberal democratic society. In G. Grant (Ed.), *Review of Research in Education: Vol. 17* (pp. 413–483). Washington, DC: AERA.

Traiman, S. L. (1993). *The debate on opportunity-to-learn standards*. Washington, DC: National Governors' Association.

Travers, K. (1993). Overview of the longitudinal version of the Second International Mathemat-

ics Study. In L. Burstein (Ed.), *The IEA study of mathematics III: Student growth and classroom processes* (pp. 1–14). Oxford: Pergamon.
U.S. Department of Commerce. (1989). Current population survey, October, 1989. Washington, DC: Bureau of the Census.
Verba, S., & Orren, G. R. (1985). *Equality in America*. Cambridge, MA: Harvard.

Authors

GRETCHEN GUITON is an Assistant Professor in the School of Education at the University of Southern California, Waite Phillips Hall, 600 D, Los Angeles, CA 90089-0031. Her area of specialization is assessment and research methods.

JEANNIE OAKES is a Professor in the Graduate School of Education and Information Studies at the University of California, Los Angeles, 2052 Moore Hall, Los Angeles, CA 90095-1521. Her area of specialization is curriculum and policy.

Received February 27, 1995
Revision received April 13, 1995
Accepted April 17, 1995