Troubling the Foundations of Special Education: Examining the Myth of the Normal Curve

Curt Dudley-Marling & Alex Gurn

One of the tasks for developing consciousness of disability issues is the attempt . . . to reverse the hegemony of the normal and to institute alternative ways of thinking about the abnormal. (Davis, 1997, p. 26)

The college professor who seeks to create a fair distribution of marks by grading “on the curve,” the Las Vegas gambler who is certain that after a series of losing rolls of the dice, his luck is sure to turn, and the special educator who determines that a child is “developmentally delayed” because she scored two standard deviations below the mean on an intelligence test—all share a tacit belief about the order of the universe. Through their respective behaviors, the professor, the gambler, and the special educator express a shared assumption about how the world works; this assumption is most commonly represented by what most of us have been taught to think of as the “normal” curve. We have all been socialized into the idea that, in the natural order of things, the achievement of students in a college course will cluster around an average grade, a very long losing streak in games of chance is rare (i.e., not normal), and certain constellations of scores on intelligence tests are exceptional.

In their controversial text, The Bell Curve, Herrnstein and Murray (1994) refer to the normal curve as “one of nature’s more remarkable uniformities” (p. 557). From this perspective, we can expect most phenomena in nature, including human behavior, to distribute normally. Indeed, it has achieved the level of common sense among educators, social scientists, and the general public that a bell-shaped
curve describes the distribution of a wide range of natural and social phenomena from income and crime statistics to various human traits (e.g., height, weight) and behaviors (e.g., intelligence, academic achievement, athletic prowess) to the “throw of the dice” (Schroeder, 2003). The conventional wisdom that the academic abilities of school children tend to cluster around “the average,” for example, stands behind the organization of schooling around age-graded curricula and whole-class instruction (Thomas & Loxley, 2001/2008). The concomitant notion that constricts human differences with reference to the average—or “normal”—is a seminal concept in the field of special education. Davis (1997) observes that

the concept of the norm . . . implies that the majority of the population must or should somehow be part of the norm. The norm pins down that majority of the population that falls under the arch of the standard bell-shaped curve.... Any bell curve will always have at its extremities those characteristics that deviate from the norm. So, with the concept of the norm comes the concept of deviations or extremes. When we think of bodies, in a society where the norm is operative, then people with disabilities will be thought of as deviants. (p. 13)

Despite its common-sense appeal, a substantial body of evidence indicates that the normal curve is a poor model of social reality that has led to “misguided educational theories, inferences, policies, and practices” (Walberg, Strykowski, Rovai, & Hung, 1984, p. 88). The normal curve adequately describes truly random events like the “throw of the dice”; however, as the Las Vegas gambler may discover, sometimes the dice are loaded (Graham, 1939). Socially mediated human behaviors, for example, do not occur randomly. Human weight, for instance, is influenced by various social, economic, and cultural factors, and, therefore, does not distribute normally (Fashing & Goertzel, 1981). Even the distribution of height among human populations may be skewed by nutritional factors. The effects of social factors on the distribution of phenomena such as wealth, academic achievement, and physical strength are even more obvious. Mistrust of the normal curve as a representation of human affairs has emerged in other fields (see Walberg et al., 1984); however, the myth of the normal curve continues to exert a powerful influence on educational thinking, particularly among special educators.

Thomas and Loxley (2001/2008) observed that “faith in certain kinds of knowledge provides the credence, the believability behind special education’s status” (p. 1). However, if one takes a “questioning disposition to this knowledge, serious challenges to the legitimacy of special education begin to emerge” (Thomas & Loxley, 2001, p. 1). The “normalization tendency in society in which given populations are viewed generally around the idea of a statistical norm” (Waterhouse, 2004, p. 72) is a foundational concept in special education. Troubling the normal curve as a representation of social realities will, therefore, have the effect of profoundly unsettling special education theory, research, and practice. At a minimum, challenging the universality of the normal curve will demand alternatives to conceptualizing individual differences in terms of deviations from a statistical mean.

In this chapter we review a body of evidence indicating that the normal curve grossly misrepresents human affairs and, therefore, is a poor model for conceptualizing human difference. We also examine the effects of questioning or troubling the assumptions of normality that are embedded in special education theory, research, and practice. Finally, we briefly consider alternative ways of conceptualizing difference. We begin by briefly reviewing the history of the concept of the normal curve.

Historical Overview of the Normal Curve

The history of the concept of “the normal curve” dates to the early 18th century when French-born mathematician Abraham de Moivre pioneered the theory of probability, formulating the mathematical formula that would later form the basis of the normal curve (Bradley, 1968). Specifically, de Moivre discovered the mathematical expression of the limiting case of a binomial distribution for chance events such as flipping a coin. A generation later, Carl Gauss and Pierre-Simon Laplace applied de Moivre’s theory to the distribution of measurement errors in astronomical observations. Gauss and Laplace determined that sightings of stars tended to bunch around the mean of the probability curve of errors, with deviations from the mean resulting from a wide array of minor, independent, and random causes. By calculating repeated observations, astronomers could use the mean to estimate the actual location of stars (Bradley, 1968).

Nineteenth-century Belgian astronomer Adolphe Quetelet appears to have been the first person to propose that the “normal curve of error” could be applied to the social realm of human beings (Hacking, 1990). Convinced that the normal curve would hold for measurements in physical and social domains, Quetelet (1842) sought to determine the average physical and behavioral characteristics of human populations through the use of descriptive statistics. Ultimately, Quetelet hoped to identify a composite of average values across multiple variables that made up the mythical “average man,” which, for him, represented an ideal of physical, behavioral, and social form. As Quetelet put it, “deviations more or less great from the mean have constituted ugliness in body as well as in morals and a state of sickness with regard to the constitution” (Quetelet, cited in Porter, 1986, p. 103). From this perspective, deviations from the mean denote errors or imperfections in design that occur in a determinate fashion that approximates a bell-shaped curve. Arguably, Quetelet’s appropriation of the normal curve as a
model for understanding variation in human behavior conflates variation with deviation and normal with natural, laying the groundwork for social Darwinism and structures of schooling that pathologize difference.

Whereas Quetelet focused on population averages, Sir Francis Galton, cousin of Charles Darwin and a founder of the eugenics movement, turned his attention to variation among human populations, in particular, deviations from the mean. For Galton, the mean represented less than the ideal since clustering around the mean were the undistinguished masses. Galton believed that the tails of the bell-shaped, normal distribution represented strength and brilliance at one end and weakness and feeble-mindedness at the other. He sought to extend Quetelet’s application of the normal curve to classify human intelligence, which Galton viewed as biologically determined (MacKenzie, 1981). In Hereditary Genius (1869), Galton argued:

This is what I am driving at—that analogy clearly shows there must be a fairly constant average mental capacity in the inhabitants of British Isles, and that deviations from that average—upwards towards genius, and downwards towards stupidity—must follow the law that governs deviations from all true averages. (Galton, 1869, p. 32, cited in MacKenzie, 1981, p. 57)

Toward this end, Galton employed the normal curve of errors to sort individuals into hierarchically based, quartile ranges based on their deviation from the mean. It was Galton who transformed the normal distribution into rankings so that one tail of the normal distribution would be seen as optimal or desirable and the other tail as abnormal and undesirable (Davis, 1997). Individuals in the lowest quartile were considered abnormal or deficient, while those in the upper quartile embodied progress and perfection. For Galton, the mean and the standard deviation provided an objective means for ranking people’s mental capacity along a continuum. Arguably, Galton’s ideas about the normal distribution stand behind the organization of American education that makes hierarchies out of “any differences that can be claimed, however falsely, to be natural, inherent, and potentially consequential in school” (McDermott, Goldman, & Varcnne, 2006, p. 12).

By the late 19th century statisticians were persuaded that natural and social phenomena would always distribute “normally” if a sufficiently large number of observations were obtained (Micceri, 1989). This assumption was challenged, however, by Karl Pearson, a pre-eminent figure in the development of modern statistics, who, on the basis of empirical observations, raised questions about the prevalence of normality among real-world distributions (Micceri, 1989). Despite Pearson’s protestations, the tendency to take for granted the normal distribution of observations in the natural and social worlds persisted, partly due to the work of eminent statistician Ronald Fisher “who showed that, when universal normality could be assumed, inferences of the widest practical usefulness could be drawn from samples of any size” (Geary, 1947, p. 241). Following Fisher’s insight, “prejudice in favor of normality returned full force” (Geary, 1947, p. 241) and challenges to the universality of the normal distribution as a representation of human variation receded to the background.

For nearly 100 years taken-for-granted assumptions of normality have undergirded educational research and test construction as scholars and administrators in the fields of education and psychology turned to positivist, quantitative science to assert the validity of their scientific claims (Lagemann, 2000). For their part, schools readily embraced the technique of measuring students against standardized norms that “involved new social practices, like standard curriculum, age grading, and examinations, which... created the kinds of statistical populations that Galtonian psychology took as its basis” (Danziger, 1990, p. 79). The concept of the normal curve has had a powerful impact on how educators and psychologists think about students, particularly students deemed to be “exceptional.” Yet, as we demonstrate below, in the social worlds inhabited by human beings normal distributions are uncommon. Further, even in instances where human behavior distributes more or less normally, the mean offers a poor representation of the behavior of individuals and groups.

The Myth of the Normal Curve

Emerging from the concept of the normal curve are two statistical measures that have shaped understandings of difference in the context of schooling: the standard deviation and the mean. The standard deviation supports the construction of human difference in terms of probability distributions based on assumptions about how various traits and abilities distribute within the general population. Intellectual and developmental disabilities, for example, have been operationally defined in terms of deviations from a statistical construction of average (i.e., the mean). In this case, the mean provides a point of reference for determining who is “normal” and who is not. The mean is also used to support the development of categories of difference. Implicit in the multitude of studies comparing categorical groupings of students with “exceptional needs” to “normal” students is the assumption that the concept of “the mean” represents both differences (from normal) and similarities (shared traits and behaviors within categorical groupings). Studies that reveal statistically significant differences in various traits and abilities between students with learning disabilities and non-disabled students, for example, support the characterization of some students as exceptional (i.e., statistically different from normal) while simultaneously validating a category of exceptionality (e.g., students with learning disabilities who share certain cognitive and behavioral characteristics).
In the following section, we critique the use of the normal curve as a representation of variation within the general population and as a way for representing specific groups and individuals.

**The normal curve as a representation of variation among the general population**

Sir Francis Galton, one of the first people to advocate the use of the normal curve as a model of human diversity, also provided one of the earliest challenges to the universality of the normal curve. Galton set out to gather a variety of empirical data in order to demonstrate the utility of the normal curve but, as it turned out, he discovered that the data for traits like height, weight, strength, and eyesight failed to produce perfect normal distributions. Ultimately, Galton concluded that the normal curve applied only to homogeneous distributions (Maccari, 1989). Galton found that he could achieve perfectly normal distributions by converting his data to standard scores and averaging these standard scores from different variables together (Fashing & Goertzel, 1981). However, according to Fashing and Goertzel (1981), the normal distributions produced from these average scores, which Galton mistook as evidence of the underlying normality of the variables, merely indicate the presence of measurement error (Fashing & Goertzel, 1981).

Karl Pearson, writing over 100 years ago, offered a more direct challenge to the universality of the normal curve. Based on his extensive observations of various phenomena (e.g., throws of the dice, outcomes of roulette wheel, number of petals in buttercups), Pearson (1900) concluded that a wide range of phenomena—many cited as textbook examples of normality—did not fit a Gaussian, or normal, curve. Pearson expressed regret that statisticians too easily accepted claims of normal distributions on theoretical and not empirical grounds. He concluded:

> If the earlier writers on probability had not proceeded so entirely from the mathematical standpoint, but had endeavored first to classify experience in deviations from the average, and then to obtain some measure of the actual goodness of fit provided by the normal curve, that curve would never have obtained its present position in the theory of errors. Even today there are those who regard it as a sort of fetish.... [However] the normal curve of error possesses no special fitness for describing errors or deviations such as arise either in observing practice or in nature. (Pearson, 1900, pp. 173–174)

David Wechsler (1935) and Lee Cronbach (1970), both major figures in the history of psychological assessment, also cautioned that psychological phenomena are not inherently distributed normally (Fashing & Goertzel, 1981). Geary (1947) went even further, recommending that all statistics textbooks begin with the statement, “Normality is a myth; there never was, and never will be, a normal distribution” (p. 241). Although Geary (1947) conceded this was an overstatement, he argued that researchers should never take normality for granted. Indeed, over time, researchers have identified numerous examples of what Bradley (1977) has called “bizarre distributions” of human behavior that differ significantly from the normal distribution. Human vigilance experiments in which the dependent variable is time to respond to critical signals, human taste thresholds, social conformity, and time to completion for various tasks are just a few examples of human behaviors that have been found to produce dramatically non-normal distributions (Bradley, 1977). In the world of human affairs, significantly non-normal distributions, often the result of uncontrolled-for variables, are quite common (Bradley, 1977). Indeed,

> there are vast areas of research (especially in the behavioral sciences) where the uncontrolled variables are likely to include unknown, or insufficiently known, variables of moderate or even major influence and, indeed, where the experimenter has practically no knowledge as to the shape of the sampled population. (Bradley, 1977, p. 150)

One type of non-normal distribution that is commonplace in the world of human affairs is the right- or positive-skew distribution—"in which low or even null performance is most frequent, and high or even moderate performance is rare" (Walberg et al., 1984, p. 87). A former president of the Royal Statistical Society declared that “positive skewness is the most pervasive law-like phenomenon in the social sciences” (Walberg et al., 1984, p. 87). Based on an extensive review of research across a variety of fields of study, Walberg et al. (1984) concluded that “positive-skew distributions characterize many fundamental processes and objects in biology, communication, crime, economics, demography, geography, industry, information and library science, linguistics, psychology, sociology, and the production and utilization of knowledge” (p. 108). This analysis supports the conclusion that the normal curve applies only to truly random events. In the natural world and in the social worlds inhabited by human beings many—and perhaps most—phenomena are not the result of random events and, therefore, do not distribute normally.

Despite widespread challenges to the normal curve as a representation of human behavior, the normal curve continues to exert a powerful influence on educators and psychometricians (Maccari, 1989). It may be that educators and psychometricians have been unduly influenced by the assumption that well-designed, objective tests necessarily produce normal distributions that are representative of human behavior. Educators may assume, for example, that learning outcomes are normally distributed because achievement scores are presumed to distribute
normally. However, achievement tests are "by tradition, custom, or conscious purpose . . . designed to produce such manifest distributions and are not necessarily indicative of the underlying latent [normal] distributions" (Walberg et al., 1984, p. 88). The second half of this quote bears repeating: Just because objective measures of achievement or ability produce normal distributions does not mean that what is being measured (math, reading, intelligence, and so on) actually distributes normally among human populations.

The tendency of achievement and ability test data to distribute normally is, to some degree, "simply a mathematical and statistical effect" (Sartori, 2006, p. 415). Standardized educational tests rely on summated scaling techniques by which persons taking tests attempt to answer a large number of items and receive total scores corresponding to the number of items they answer correctly. This type of measurement has an inherent bias towards a normal distribution in that it is essentially an averaging process, and the central limit theorem shows that distributions of means tend to be normally distributed (Fashing & Goertzel, 1981; Sartori, 2006). In other words, the average of averages will necessarily produce a normal distribution even if the variable being measured does not distribute normally.

The tendency of objective tests to produce normal distributions is also a function of item difficulty (Walberg et al., 1984) and test error. Recall that the normal curve was originally referred to either as the "law of frequency of error" or "the normal curve of error" (Hacking, 1990; MacKenzie, 1981). Sartori (2006) observes:

When the responses to the items of a test or scale are poorly intercorrelated and when a large number of people fill it out, the scores are very likely to be normally distributed. This characteristic of the averaging process is useful in calculating probable errors in random sampling. But when averaging is used in testing or measurement, it may mean that the greater the amount of error, the greater the likelihood of a normal distribution of scores, even if . . . the variable being measured is not normally distributed. All objective tests contain a certain amount of error . . . Thus, it is not surprising that summated (or averaged) scaling devices tend to give normal distributions. (p. 415)

Crucially, "the problem comes when this effect is interpreted not as the obvious result of an error which is unavoidable, but as a confirmation of a preconceived idea that the variables being measured are really normally distributed" (Sartori, 2006, p. 415).

Even given the theoretical bias of objective tests toward normal distributions there is empirical evidence indicating that actual test scores "are seldom normally distributed" (Nunnally, 1978, p. 160). Micceri (1989), for example, examined the distributional characteristics of 440 large-sample achievement and psychometric measures obtained from journal articles, research studies and national, state, and district tests. Major sources of test data included the California Achievement Tests, the Comprehensive Test of Basic Skills, Stanford Reading Tests, Scholastic Aptitude Test (SAT), and the Graduate Record Exam (GRE). In all, Micceri's sample included 46 different test sources and 89 different populations. His analysis indicated that all of the 440 distributions he examined were "significantly non-normal at the alpha .01 significance level" (p. 156) reinforcing Karl Pearson's conclusion that "the normal curve of error possesses no special fitness for describing errors or deviations such as arise either in observing practice or in nature" (Pearson, 1900, p. 174), including the distribution of test scores.

The presumption that social and psychological variables distribute normally "has been shown to be invalid by those methodologists who have taken the trouble to check it out. Its persistence in the folklore and procedures of social institutions is a reflection of institutionalized bias, not scientific rigor" (Fashing & Goertzel, 1981, p. 28). The normal curve is a product of random errors. Human behavior, which is always influenced by social factors, is never random and, therefore, should not be expected to produce normal distributions (Fashing & Goertzel, 1981). Still, among educators and psychometricians, the conundrum of the normal curve has largely been treated as a technical problem, not a conceptual one. Many educational researchers, for example, appear to take comfort in Monte Carlo studies indicating the robustness of parametric statistical tests in the face of non-normal distributions, yet ignore what the normal curve signifies: a theory of human difference that has no empirical basis.

The normal curve as a representation of individuals: The myth of the "average animal"

The myth of the normal curve has given rise to the premise that the mean is a meaningful representation of groups—and individual group members—to which it is applied. In the theoretical case of a perfect normal distribution, observations will cluster about the mean. It is well known that, in the case of phenomena that result in a perfect, bell-shaped curve, 95% of observations will fall within two standard deviations of the mean. Sixty-eight percent fall within one standard deviation of the mean. When the normal curve is applied to human populations, the area demarcated by one standard deviation on either side of the mean is often referred to as the "average range," a statistically constructed space where people are assumed to share essential characteristics by virtue of the simple fact that they do not differ significantly from the norm (i.e., average). Reading achievement tests, for example, construct "average readers" who, presumably, share essential learning traits and require common curricula. These same tests also play a role in defining students who fall outside the average range, students whose profiles
differ from average—or normal—and who, it is assumed, share essential qualities that may require special education.

The mean as a representation of average (or normal) within a normal distribution is fundamental to the special education enterprise. The mean constructs categories of exceptionality by identifying traits and abilities that separate children with disabilities from the general population of students (e.g., students with intellectual disabilities are defined as performing significantly below the mean on intelligence tests) providing the primary rationale for special education. The use of means to indicate traits and abilities groups of children with disabilities share (e.g., the average student with learning disabilities) justifies and maintains special education research and practice focused on categories of exceptionality. Research into best instructional practices for students with learning disabilities, for example, presumes that students with learning disabilities share characteristics related to their learning that distinguish them from students who are not learning disabled.

The problem is that the normal curve is a poor representation of human differences. As the evidence we reviewed above indicates, human traits do not tend to cluster about the mean in a bell-shaped distribution. Yet, even in the theoretical case of a perfect normal distribution, the mean provides a misleading portrayal of individual traits and abilities within particular populations. Writing over 70 years ago in the *Journal of Comparative Psychology*, Knight Dunlap (1935) warned of the dangers of reporting data solely in terms of the average animal, “an animal which is entirely mythical” (p. 1). Dunlap observed that in his “list of Great Experiments in Bad Psychology there is one research in which the average value presented as significant is a value which every person in the experiment conspicuously avoided” (p. 2). Put differently, the average for any particular group of people may apply to no one person in the group. In the context of special education research, the reliance on means to represent groups of students with disabilities mischaracterizes individual students. The finding that the average student with learning disabilities, for example, is deficient in some trait compared to non-learning disabled students may not apply to individual students with learning disabilities. Dudley-Marling, Kaufman, and Tarver (1981), for example, reported that learning profiles that were claimed to be characteristic of students with learning disabilities were not always found in individual students with learning disabilities. Additionally, it was found that many non-LD students did present these profiles. Similarly, demonstrating the effectiveness of an instructional strategy with a statistically constructed average student with intellectual disabilities obscures the likelihood that the strategy was ineffective with at least some children. No reading intervention, for example, has been found to be successful with all children, all of the time (Duffy & Hoffman, 1999) even if various interventions have been found to be effective on average.

Overall, the statistically constructed average student with disabilities is a mythical individual who bears little resemblance to individual students with disabilities. Yet, statistical averages have frequently been the basis of generalizations about students with disabilities “even though the average pattern might not correspond to a single individual member of a statistical group” (Danziger, 1990, p. 153). The irony is that special education, which focuses on the needs of individual students, is undermined by a model of human variation that tends to efface individual differences.

**Conclusion**

The concept of normality, in which human diversity is characterized by a bell-shaped curve and “given populations are viewed generally around the idea of a statistical norm” (Waterhouse, 2004, p. 72), is “one of the most powerful ideological tools” of modern society (Hacking, 1990, p. 169). Davis (1997) observed that “there is probably no area of contemporary life in which some idea of a norm, mean, or average has not been calculated” (p. 9). The assumption, based on the normal curve, that it is normal — and natural—for some proportion of the population to be well below average has been used to justify persistent social, economic, vocational, and academic inequities that plague contemporary American society. There is no reason to pursue public policies aimed at eliminating poverty, for example, if extremes in income distribution—and competence—are viewed as natural manifestations of the normal curve (e.g., Herrnstein & Murray, 1994). As Fashing and Goertzel (1981) put it,

> the myth of the normal curve has occupied a central place in the theory of inequality. Apologists for inequality in all spheres of social life have used the theory of the normal curve, explicitly and implicitly, in developing moral rationalizations to justify the status quo. (p. 15)

In the context of schooling, the assumption that the academic performance of similarly aged students tends to cluster about the mean justifies whole class instruction and one-size-fits-all curricula that efface individual differences among students. Indeed, the inability of some students to profit from routine, whole class instruction is taken as *prima facie* evidence that they fall outside the normal range. Further, taking the normal curve as a model of human variation suggests that this situation is natural. The theory of the normal curve predicts that a fraction of students will necessarily fall at the lower end of the normal distribution. These students are deemed to be exceptional. In this way, the normal curve provides a rationale for special education that situates learning problems in the heads of individual learners and not in the structures of schooling that produced so
much failure in the first place (Dudley-Marling, 2004). The theory of the normal curve also underpins an approach to special education research that relies on studies of mean differences to establish categories of exceptionality based on shared characteristics that distinguish students in one category of exceptionality from students in other special education categories and from normally achieving students. Studies of mean differences are also used to identify best practices for students in particular special education categories.

We argued at the beginning of this chapter that special education theory, research, and practice are underpinned by a conceptualization of normality based on the normal curve. However, as this review has illustrated, the normal curve is "normal" only in the case of random errors. There is "no reason to expect sociological variables to be normally distributed. Nor is there any reason to expect psychological variables to be if they are influenced by social factors" (Fashing & Goertzel, 1981, p. 27) as they always are. The normal curve is a poor model of social reality, and, therefore, human diversity cannot be understood with reference to a fictitious average (or normal) person. Nor can individual human differences be relegated to the distant boundaries of a bell-shaped distribution.

Exploding the myth of the normal curve undermines the representation of disability (or difference) as deviation from the mythical "normal." Consequently, we need alternative ways of conceptualizing the meaning of difference that is not based on the normal curve of errors. In the opening of this chapter we quoted Lennard Davis (1997) who issued a clarion call to disability studies theorists "to reverse the hegemony of the normal and to institute alternative ways of thinking about the abnormal" (p. 26). Upsetting the normal curve as a model of human diversity challenges the practice of equating difference with deviance and diversity with exceptional. In the realm of human experience, diversity is not the exception but the rule (Gould, 1997). In other words, it is normal to be different, a conclusion that demands a radically different conceptualization of special education theory, research, and practice.

Interrogating the normal curve as a foundational concept in special education reveals a fundamental anomaly in special education theory, research, and practice. Individualization is the raison d’être of special education; yet, special education theory, research, and practice are informed by a model of diversity that obscures individual differences. From the perspective of individual students, the discourse of science, as taken up in special education, obfuscates more than it enlightens (Gallagher, 1998; Gladwell, 2006). Relying on aggregate measures of central tendency, for example, provides a means for special education researchers to cope with diverse and inconsistent behaviors of individuals (Danziger, 1990). However, "the reference to an aggregate seems to be an unnecessary extra step if the goal of the particular study is to understand how development takes place within a given organism or psychological system..." [Further] aggregation of data entails ignoring the holistic nature of the phenomena from which those data were derived" (Surjan, 2001).

If it is normal to be different, "it is only by attending with care to each child that the noble aim of equality of education for all children can be achieved" (Carini & Himley, 2010). "Attending with care to each child" requires approaches to research—case studies, single subject research, classroom ethnography, and teacher research, for example—that illuminate the development of individual students. At a minimum, special education researchers—indeed, all educational researchers—must give more attention to reporting measures of variance in their data. In intervention studies, for example, reporting the proportion of students in treatment groups that actually improve relative to control groups (and the proportion of students in control groups who outperformed students in the treatment groups) should be a routine practice. Similarly, researchers who report effect sizes should also indicate the degree of overlap in the distributions of scores for treatment and control groups. For example, an effect size of .8, which in the world of social science is considered to be "large" (Cohen, 1969), means that 79% of the control group falls below the mean for the experimental or treatment group. It also means, however, that 21% of the control group scored higher than the mean for the treatment group (Coie, 2002). Of course, the meaning of effect sizes is much more difficult to ascertain in the case of significantly non-normal distributions which, as we have argued, are common in human affairs.

Attending to the needs of individual children also demands instruction that responds to the development, background knowledge, and experience of particular children. Curriculum-based instruction, universal design, and response-to-intervention are promising developments in this direction. Equally important is creating classroom structures that enable teachers to provide students with frequent, intensive, explicit, and individualized support and direction based on ongoing, individualized assessment (Dudley-Marling & Paugh, 2004). Readers and Writers Workshops and center-based instruction are examples of classroom organization that provide opportunities for teachers to work with students individually and in small groups and collect the in-depth assessment data needed to provide appropriate instruction for individual students.

The normal curve is a powerful construct that, in the context of schooling, limits the interpretations available for thinking and talking about children (McDermott, Goldman, & Varèine, 2006). 'Teachers' interpretations of student learning typically involve categorizing each child in relation to a significant social boundary—based on the normal curve—"that separates 'normal' from 'deviant' pupils" (Warhouse, 2004, p. 72). So challenging the normal curve means confronting the everyday language educators use to interpret students' learning. As a beginning we need to challenge teachers, researchers, and school administrators to focus less on what children cannot do and more on what they are capable
of given the right conditions, including challenging curriculum. As McDermott, Goldman, and Varènne (2006) put it, “to counteract the cultural inclination to focus on what is wrong with individual children, we must seek data showing children more skilled than schools have categories or time to notice, describe, diagnose, record, and remediate” (p. 15). We must stop asking what’s wrong with children who struggle in school—that is, how they compare to some mythical norm—and ask, instead, what makes them smart (Miller, 1993).

Endnotes
1. Standard scores are also called z scores which indicate distance from the mean in standard deviation units.
2. Skewness indicates the extent to which a distribution is asymmetrical. In a right- or positive-skew distribution the tail of the distribution extends to the right. Income among American families is an example of a positive-skew distribution, i.e., most people make less than $100,000 and only a small proportion make more than $1,000,000. In other words, income distributions are skewed toward lower incomes.

References