The Sociopsychometrics of Learning Disabilities

Keith E. Stanovich

Abstract

The Boston University (BU) case illustrates how the psychometrics of ability differences interact with the concept of learning disability and with the sociopolitics of schooling and society. It also illustrates that learning disabilities advocacy will not be on a sound footing as long as the field refuses to rid itself of its IQ fetishism, refuses to jettison its fixation on aptitude—achievement discrepancy, and fails to free clinical practice from the pseudoscientific neurology that plagued the field in the 1970s. A more inclusive definition of learning disability—one that abandons discrepancy notions—and a more self-critical attitude toward its own claims would advance the field of learning disabilities and help to rid it of distractions such as the BU case.

After reading through the entire 111-page Saris decision in the Guckenberger et al. v. Boston University case and the comments in this issue, one cannot help thinking that President Jon Westling of Boston University (BU) and the field of learning disabilities (LD) deserve each other. Both appear to lack self-critical tendencies and to combine this trait with a moral righteousness that is particularly grating. It seems inevitable that the LD field’s sloppiness, its excesses, and its unwillingness to fully examine its social implications would spawn public criticism of this type (see Metcalf, 1998; Shalit, 1997). It seems equally inevitable that President Westling’s ill-informed remarks and behavior would spawn a lawsuit and hamper Boston University’s position once the litigation had begun.

I will focus in this commentary on how the Boston University case highlights certain lessons that the LD field has failed to learn. This commentary will not contain my usual extensive reference list because the empirical base for my position has been extensively documented in my previous critiques of the LD field (Stanovich, 1986, 1991, 1993, 1994, 1996b) and, of course, in the writings of numerous other sympathetic critics such as Linda Siegel, Jack Fletcher, Joe Torgesen, Reid Lyon, and many others. In what follows, I will outline the conceptual warnings to the field that are present in the six commentaries in this issue and in the controversy surrounding the Boston University case.

Conceptual Confusion

The term learning disability is redundant and semantically confusing. If the term is taken literally, it is not needed. We already have terms for generic learning difficulties. These terms are mental retardation, low intelligence, slow learner, and so forth. The field, of course, does not take the term literally but uses it in a counter-intuitive way, to apply to a range of domain-specific problems (e.g., reading, arithmetic) that simply do not form a natural kind (a situation exacerbated when attention-deficit/hyperactivity disorder and social skill deficits are thrown into the mixture, as they sometimes are).

The umbrella term learning disabilities only causes confusion. The domain-specific disabilities should be treated separately and be labeled separately (reading disability, arithmetic disability, etc.). Comorbidity becomes an issue after the initial domain-specific classification has been carried out. Both for purposes of service delivery and for scientific purposes, the term LD does nothing but confuse. The issue of accommodations—one of the foci of the BU case—is likewise confused by the term LD. Accommodations, if deemed appropriate after a social policy debate (a debate barely begun; see Kelman & Lester, 1997), should be for particular, domain-specific disabilities.

Reading Disability as a Model

We know more about reading disability than about all the other so-called learning disabilities put together. If the field cannot agree on rational criteria for this particular disability, it is futile to try to establish disabilities in domains that are less fully researched and less domain specific. If we cannot get reading disability correct (i.e., work out agreed-upon classification criteria and agreed-upon accommodations without untoward unintended consequences), at-
tempting to deal with other purported disabilities would represent little more than social policy based on pseudo-science. The field’s ability to deal rationally with reading disability would be prima facie evidence of its ability to deal with other, more esoteric disabilities.

Siegel (in this issue) has outlined what is known about reading disability. Its proximal cause is difficulties with word recognition skills due to weak grapheme-to-phoneme coding skills. These skills are weak because of segmental language difficulties (lack of phonological awareness). The phonological deficit fulfills what has been termed (see Stanovich, 1986) the assumption of specificity in reading disabilities research—the assumption that the underlying deficit must be dissociable from intelligence. This assumption underlies all discussions of the concept of reading disability (traditionally defined), even if it is not explicitly stated. It is the idea that an individual with this type of LD has a brain–cognitive deficit that is reasonably specific to the reading task. That is, the concept of reading disability (or dyslexia) requires that the deficits displayed by such individuals not extend too far into other domains of cognitive functioning. If they did, this would depress the constellation of abilities we call intelligence and reduce the reading–intelligence discrepancy central to all current definitions (see Siegel, 1988, 1989; Stanovich, 1986, 1991). In short, the key deficit in reading disability must be a domain-specific process (see Fodor, 1983) rather than a central cognitive mechanism with widely distributed effects (what Fodor calls central processes). A candidate for the processing deficit underlying reading disability must meet the specificity assumption if reading disabilities are to occur throughout the intelligence continuum rather than be a property solely of individuals with low IQ. Phonological processing and phonological awareness meet this criterion of partial modularity (Fodor, 1983; Stanovich, 1986, 1990, 1996b).

Does the field in practice reflect this knowledge about reading disability—the most common of the so-called learning disabilities? From the documentation submitted to Boston University and examined by Siegel (in this issue), the answer is clearly no. Siegel reports the existence in these files of a depressingly high number of vague and empirically unverified terms from the prescientific age of LD research (visual processing speed, visuo-perceptual processing deficits, ability to process auditory and visual information, etc.)—many such terms reflecting central rather than modular processes (see Fodor, 1983) and violating the specificity assumption that must be central to any domain-specific disability purported to dissociate from intelligence.

Siegel (in this issue) also reports that, in a number of cases, no achievement test results were in the students’ records, thus precluding the valid identification of the student as having the most common and extensively studied learning disability. Siegel rightly notes that “any assessment designed to detect the presence of dyslexia should include an assessment of phonological skills” and points out that such an assessment was missing from many files. I endorse her call for the performance on a test of pseudoword reading (e.g., Woodcock Word Attack subtest) to be the primary diagnostic indicator of reading disability. The most common and accurately identified type of reading disability is phonological dyslexia (see Castles & Coltheart, 1993; Manis, Seidenberg, Doi, McBride-Chang, & Peterson, 1996; Stanovich, Siegel, & Gottardo, 1997), and its primary indicator is difficulty in pronouncing pseudowords. However, in order to pick up the numerically smaller and less easily identified surface subtype (see Castles & Coltheart, 1993; Stanovich et al., 1997), a test of word recognition (e.g., Woodcock Word Identification, Wide Range Achievement Test) can be used in conjunction. If discrete classification is a necessity, then in order to identify both subtypes I would recommend that an “or” criterion be used with these two measures—performance on either of these two tests below a certain percentile cutoff would define a reading disability. Intelligence would play no role in the diagnosis.

Because this classification would be made by the disjunction of two measurements, the 25th-percentile criterion discussed in Siegel’s commentary (in this issue) would likely be too liberal. I would probably opt for a more stringent criterion such as the 15th percentile, or even the 10th, on at least one of the tests. However, this is a subjective and arbitrary social policy recommendation (because the relevant distributions are continuous rather than discrete) rather than a purely scientific one (more on this further on). It should be made after an explicit discussion of competing resource demands on the relevant educational system. I base my own guesstimate on inferences from recent training studies that have examined the difficulty of remediation (Blachman, 1997; Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Lovett et al., 1994; Olson, Wise, Johnson, & Ring, 1997; Stanovich & Stanovich, 1997; Torgesen, 1997; Torgesen, Wagner, & Rashotte, 1997; Vellutino et al., 1996). Because of the continuous nature of the distribution of reading ability, no percentile cutoff should be retied in legislative statutes. Instead, this cutoff must be calibrated by universities and school districts based on social policy decisions that take into account competing demands on resources.

I would note in passing the interesting fact that my recommendation for the assessment information needed to classify a person as having a reading disability represents the exact converse of the contents of some of the files that Siegel examined. Some of these files contained results of the administration of intelligence and neuropsychological tests, but no results at all from scholastic achievement tests. My proposal (and Siegel’s) for identi-
flying reading disability necessitates the administration of an achievement measure (e.g., the Woodcock) but not of intelligence or neuropsychological measures.

Siegel’s description of these case files from Boston University illustrates why the university’s demand that the students present recent assessments in order to qualify for special accommodations (extra time on tests, etc.) was well founded. The debate on this point in the actual litigation apparently focused on whether reading disability is a transitory problem or a long-lasting developmental disability. This, I submit, is not the point at all. The issue is not whether reading disability changes; the issue is changes in the field! Standards of clinical practice in the LD field have, at various times in its history, reflected more than 50% pseudoscience or unverified, virtually armchair speculation about human abilities. Such unverified knowledge has seeped into professional practice and is difficult to dislodge. A recent assessment (and probably a PhD assessor, as Siegel and BU argued) provides the only hope that the assessor will be passably familiar with current knowledge on reading disability and its phono logical basis. Oft assessments by ill-trained professionals may yield the litany of vague, nonmodular processing “diagnoses” that Siegel mentions in her commentary (see Note 1).

University policymakers concerned with accommodations should not be forced to ratify the field’s historic mistakes because the LD field has introduced psychometric and theoretical errors into law and professional practice.

Psychometric and Theoretical Errors

Imagine that your current HMO covered only those procedures and diseases recognized by the medical profession in 1950. The thought would be ridiculous, because in such a rapidly developing field as medicine no one would expect practice to be frozen at the level of scientific knowledge attained 50 years ago. In a less extreme fashion, though, just this has happened in the LD field. The case documentation that Siegel (in this issue) comments on seems clearly reflective of the state of the LD field in the 1970s. Advocates and special-interest groups such as the Boston University plaintiffs are wrong to try to freeze the field at a point when our state of knowledge was archaic (and are doubly wrong to try to freeze practice by legislative mandate and litigation). The LD field suffers greatly from its tendency to base practice on concepts and psychometric technologies superseded by subsequent scientific advances. In this section, I will mention several examples (see Fletcher, Foorman, Shaywitz, & Shaywitz, in press; Fletcher et al., 1994; Kavale & Forness, 1995; Siegel, 1989, 1993; Stanovich, 1986, 1991, 1996b; Stanovich & Siegel, 1994; for more extensive discussions).

Discrepancy Definitions Are Untenable

Practice in the LD field has yet to catch up with the fact that a decade’s worth of research has undermined the rationale for defining reading disability by reference to aptitude—achievement discrepancies. A host of recent studies have demonstrated that the information-processing operations that underlie the word recognition deficits of poor readers are the same for poor readers with low IQ and high IQ (Fletcher et al., in press; Fletcher et al., 1994; Siegel, 1988, 1989; Stanovich & Siegel, 1994). There is no evidence that low-IQ and high-IQ poor readers respond differently to treatment. There is also no evidence that the neuroanatomical defects that underlie the cognitive deficits of these two groups are different.

The excessive focus on intelligence in the LD field (“IQ fetishism”—to be discussed further on) stems from two sources, one social and the other conceptual (hence my reference to the sociopsychometrics of LD). The conceptual error (see Stanovich, 1991) derives from the early focus of the LD field on the notion of “unexplained” reading failure. This layperson’s notion of unexplained reading failure—a type of reading failure somehow different from reading failure having an explanation—turns out to be nothing but muddled folk psychology (Churchland, 1989) that scientific progress in the reading disability field is in the process of replacing.

Intelligence has played a major role in the conceptual muddle surrounding the notion of unexplained reading failure. The confusion arises because it makes no sense to say that low intelligence (intelligence being defined as the statistical amalgamation of a panoply of different cognitive processes) causes reading difficulties, given what is currently known about reading disability. To assume so would be to make the fundamental reasoning error of inferring cause from (possibly spurious) correlation—to assume that everything bad that is correlated with low IQ is caused by low IQ. The concept of intelligence does not provide the specific model that explains poor reading. The phonological deficit model does (see Share & Stanovich, 1995, and Siegel, in this issue). It would be perfectly reasonable to inquire whether the processing mechanism that is the cause of reading difficulties is correlated with IQ. However, if there were a correlation, it would make low IQ merely a marker variable, not a cause of poor reading.

An Illustrative Example. The following example illustrates the point. Every year, many people die from structural heart problems. Many of these people are not very physically fit. However, once every couple of years, a highly trained professional athlete keels over and dies of a structural heart problem during a professional athletic contest. The athlete experiences “unexpected heart failure.” As this heart failure was so un-
expected (we don’t expect it in people who are physically fit), should we automatically assume that it must have a cause different from the structural heart problems of individuals who are not so fit? Clearly not. We cannot assume that lack of fitness had anything to do with the structural problem, and we equally cannot assume that the unexpectedness of the athlete’s death points to a different cause. Everyone’s heart failure has the same cause—some people who die just happened to be fit and some were not. The athlete’s death is only unexpected because the physiology of the individual was opaque to us. Had we known the relevant physiological facts, the athlete’s death would have been expected. It is only our ignorance of the physiological mechanism—and the natural expectation that good physical things go together just as good cognitive things do—that made the death unexpected.

Likewise, we can measure phonological processes in children, and we can predict who will fail to acquire reading easily (just as we could have predicted the heart failure from the relevant physical facts). The reading failure of an individual with a high IQ is expected if this person has a deficit in phonological awareness. Only ignorance of current models of reading difficulty and of theories of intelligence lead a layperson to consider reading failure in an individual with a high IQ to be unexpected. It is actually perfectly expected given our current knowledge of the modular nature of many information processing skills (Carr, 1992; Perfetti & McCutchen, 1987; Posner & Carr, 1992; Stanovich, 1990). Note that, consistent with this argument, when very young readers are given a test of phonological awareness and a test of intelligence, the former is a better predictor of their subsequent reading achievement (see Stanovich, 1992, for a review of 14 different studies where this finding was replicated).

The poor reading of individuals with low IQ is not explained by their low IQ. We need a specific processing explanation for their poor reading, just as we need a specific processing explanation for the poor reading of high-IQ individuals. Whether the processing explanation will be the same for readers with high IQs and low IQs is an empirical question. Empirical evidence cited previously indicates that the processing model is the same. To phrase the conclusion in a form that highlights the appropriate null hypothesis, there is no converging empirical evidence that the processing mechanism accounting for the primary word recognition problems of poor readers with high IQs is different from the processing mechanism accounting for the primary word recognition problems of poor readers with low IQs.

An Imaginary Lawsuit. I have mentioned previously that I think that the attempts by the BU plaintiffs and other advocates to freeze practice by litigation is a mistake and a disservice to the field. Litigation will backfire—will prevent legislative attempts—and it will deserve to backfire. I have tried in the past to teach the LD field to honestly address its conceptual problems and theoretical confusions by proposing thought experiments (Stanovich, 1993). Here is another thought experiment for those in the LD field enamored with the idea of shaping practice by litigation rather than by research. If the LD field persists in its fixation on intelligence and aptitude—achievement discrepancy, for those who love lawsuits there is an obvious one just waiting to happen. My imaginary lawsuit involves a university student (presumably at a nonselective institution) with below-average reading ability but with no discrepancy from intelligence and no LD diagnosis (and hence no accommodations). It would be an easy thing for that student’s advocate to find other students with equivalent reading scores but with the discrepancy-triggered LD diagnosis—and hence with a host of accommodations, including extra time on exams. In such a case, as Siegel (in this issue) notes, the use of the intelligence test (as opposed to a pure achievement cutoff) would be the direct cause of discriminatory treatment that would be indefensible from any standpoint. Now that would be a lawsuit that the LD field deserves!

IQ Fetishism in the LD Field

The focus on IQ—achievement discrepancy in the LD field despite lack of empirical support for this fixation has been termed IQ fetishism by Kelman and Lester (1997), who argued that “while accommodation advocates tend not to draw attention to their proclivity to deify something like ‘pure intelligence,’ this is precisely what they appear to do. . . . [because they endorse the theory] that society is especially obligated to try to ensure that all students develop skills congruent with the expectations we derive from their IQ tests” (p. 175).

Kelman and Lester (1997) pointed out that it is rare for the advocates of discrepancy-based definitions to articulate the theory of social justice that dictates that society has a special obligation to bring up the achievements of individuals whose achievements fall short of their IQs, rather than simply to bring up the skills of those with low skills, period. It is in fact quite rare to see an explicit defense of this position by advocates of traditional discrepancy-based definitions (of the notion of unexpected reading failure, which amounts to the same thing). The absence of an explicit argument about social justice considerations serves to conceal assumptions about IQ test scores as measures of cognitive potential. Also concealed are assumptions about how educational resources should be allocated.

Kelman and Lester (1997) eloquently exposed these hidden assumptions with observations such as,

The idea that one’s IQ is one’s basic defining virtue—and that it is somehow "more real" or more telling a description of some “true self” than one’s in-
ability to read or pay attention—would seem persuasive only to those who have fetishized the IQ score for other reasons. Thus the idea that tests fail to reward virtue if they fail to reward high IQ would seem to be especially attractive to those loath to treat any members of the social groups that traditionally score well on the IQ test as incapable in the “same ways” as members of socially subordinated groups are. (pp. 175-176)

There is in fact very little to differentiate the assumptions about intelligence implicit in the LD field from the assumptions about intelligence made by Jensen (1998) or by Herrnstein and Murray (1994). This connection is rarely made by advocates, for fear of scaring off the liberal allies that are prone to rally around any civil rights claim, regardless of how philosophically tenuous. Kelman and Lester (1997) exposed how the LD field has used IQ tests to promulgate extreme (and often odd) notions of intellectual potential:

The metaphor that animates the IQ fetishist is the (unself-consciously paradoxical) idea of “unrealizable potential,” a core of essential personal identity that transcends behavior revealed in any actual setting. Having put so much pressure on IQ to do the serious work of social legitimation, fetishists come to believe that differences in IQ are the fundamental, socially pivotal differences among people. At that point, one’s IQ becomes synonymous with one’s basic nature, and a disparity between IQ and performance comes to indicate a breach in the natural order, a failure of one’s true nature to be manifest in practice. (p. 134, p. 137)

I have made the same argument (Stanovich, 1991) in a way perhaps slightly more relevant to the specific educational context that is the focus of this special series. I argued that one reason that discrepancy definitions so quickly became the defining feature of the LD concept was the widespread acceptance of the belief that IQ scores were valid measures of intellectual potential. Certainly an extreme form of this belief can be seen in the promotional activities of many advocacy groups and in media portrayals. The typical “media dyslexic” is almost always a very bright child who is deeply troubled in school because a glitch (usually assumed to be biologically based; see Coles, 1987) prevents him or her from reading. The subtext of the portrayal clearly implies that the tragedy of the situation is proportionally greater because the child’s great potential remains unrealized. This media portrayal has now entered the realm of folk belief, for there exists a popular myth that dyslexia is the “affliction of geniuses” (Coles, 1987)—if anything, more likely to occur in very bright people. This popular belief in the idea of unrealized potential undoubtedly helped to fuel the rapid expansion of the LD field.

One major problem, however, is that most psychometricians, developmental psychologists, and educational psychologists long ago gave up the belief that IQ test scores measures potential in any valid sense. Indeed, standard texts in educational measurement and assessment routinely warn against interpreting IQ scores as measures of intellectual potential. At best, IQ test scores are gross measures of current cognitive functioning. In short, we have been basing systems of educational classification in the area of reading disabilities on special claims of unique potential that are neither conceptually nor psychometrically justifiable.

The choice of IQ test performance as the baseline from which to measure achievement discrepancies was accepted by teachers, schools, professional organizations, and government agencies in the absence of much critical discussion or research evidence. Indeed, it is surprising that for so long the concept of intelligence received so little discussion in the LD literature. Researchers and practitioners in the field seem not to have realized that it is a foundational concept for the very idea of dyslexia. As currently defined, IQ is a superordinate construct for the classification of a child as having a reading disability. But consider the fact that researchers, let alone practitioners, cannot agree on the type of IQ score that should be used in the measurement of discrepancy. For example, it has often been pointed out that changes in the characteristics of the IQ test used will result in different subgroups of children being identified as discrepant and alter the types of processing deficits that they will display in comparison studies (see Stanovich, 1991). Yet it is not hard to look in the research literature and find recommendations that are all over the map.

For example, a very common recommendation in the LD literature is that performance or nonverbal IQ tests be used to assess discrepancy (e.g., Beech & Harding, 1984; Thomson, 1982), because verbally loaded measures are allegedly unfair to dyslexic children. In complete contrast, in an issue of Learning Disabilities Research devoted to the issue of measuring severe discrepancy, Hessler (1987) argued for the use of verbally loaded tests, because “using a nonverbal test of intelligence because an individual has better nonverbal cognitive abilities than verbal cognitive abilities does not, of course, remove the importance of verbal processing and knowledge structures in academic achievement; it only obscures their importance and perhaps provides unrealistic expectations for an individual’s academic achievement” (p. 46).

It is rarely noted that the use of certain types of intelligence tests in the operationalization of dyslexia often conceals conceptions of potential that are questionable, if not downright illogical. Consider some of the hidden assumptions behind the often-heard admonition that verbally loaded intelligence tests are unfair to dyslexic children and that performance IQ measures provide fairer measures of the reading potential of such children. Typical arguments are that “the instrument [WISC-R] is confounded and not a true measure of potential. The
learning disability itself is reflected clearly in the IQ performance" (Birnbaum, 1990, p. 330), and that “computing an IQ from items shown to be specifically associated with dyslexic difficulties may be an underestimate” (Thomson, 1982, p. 94). But it is not at all clear—even if one accepts the problematic notion of educational potential—that the spatial abilities, fluid intelligence, and problem-solving abilities tapped by most performance tests provide the best measures of the potential to comprehend verbal material. On the contrary, it would appear that verbally loaded measures would provide the best estimates of how much a poor reader could get from written text if their deficient decoding skills were to be remediated.

It goes largely unnoticed that, when people make the “fairness” argument for the use of nonverbal tests, they jettison the notion of “potential,” at least in its common meaning. They cannot mean the potential for verbal comprehension through print if the decoding deficit were remediated, because this is not what IQ tests—particularly the performance tests they are recommending—assess. What people who make the “verbal IQ scores are unfair to dyslexics” argument are asserting implicitly is that if we had a society that was not so organized around literacy, dyslexics would have the potential to do much better. True—but this is a counterfactual premise that contradicts more common usages, to the advantage of some and to the disadvantage of others. It simply makes little sense to adopt linguistic usages of the term “potential” that require the assumption that literacy-based technological societies will be totally reconstructed.

We seem to find it difficult to use this crude cognitive probe, an IQ score, as a circumscribed behavioral index without loading social, and indeed metaphysical (Scheffler, 1985), baggage onto it. If these tests are mere predictors of school performance, let’s treat them as such. If we do, performance IQ is manifestly not the predictor that we want to use, at least in the domain of literacy-based educational prediction.

Reading difficulty is a serious problem—increasingly so in a competitive economy where those lacking high-level literacy skills are disadvantaged. People with reading problems (arising from poor home environments, poor schooling, neurological insult, or whatever) need help, and society should be willing to tax itself enough that sufficient resources are provided for all to get help. None of the critics of discrepancy definitions are denying the existence of severe reading disability per se or the importance of remedial help. Instead, they are questioning the rationale of differential treatment and resources being allocated on the basis of IQ-achievement discrepancy.

**Standard Score Discrepancy Problems**

There is another mistaken practice in the LD field that interacts in a particularly striking way with the mistaken focus on discrepancy. These two interacting mistakes form part of the tangle of statistical confusions and socioeducational politics that I have chosen to call the sociopsychometrics of learning disability.

If a student labeled as having LD presented assessment results in which her or his raw score on the Woodcock Reading Mastery Tests was miscalculated and thus her or his derived scores (standard scores, grade equivalents, percentile ranks) were miscalculated (e.g., a percentile rank of 13 rather than 47 was entered), we would say that an arithmetic or recording error had taken place. If a student had been classified as LD because of such an error, we would be justified in saying that she or he had acquired her diagnosis by mistake. What goes unremarked in the LD field is that many students acquire their LD labels by a statistical mistake only slightly more complex than this.

One of the most pernicious practices in our field that still persists is that many clinicians and diagnosticians (as well as many state legislative guidelines) continue to define aptitude-achievement discrepancies in terms of differences in *standard scores* rather than in terms of regression calculations that appropriately adjust for the imperfect correlation between achievement and intelligence (Frankenberger & Fronzaglio, 1991; Mercer, Jordan, Allsopp, & Mercer, 1996). As has been much commented on (Fletcher et al., 1994; Kavale & Forness, 1995; Kelman & Lester, 1997; Reynolds, 1985; Shepard, 1980), it is easy to demonstrate that the standard score discrepancy method overidentifies individuals with high IQs and underidentifies individuals with low IQs. The reason for this is that the standard score discrepancy method does not take into account the less-than-perfect correlation between intelligence and achievement (i.e., it fails to account for statistical regression) and thus overpredicts the reading performance of individuals with high IQs.

The overidentification of children with high IQs by the standard score discrepancy method is so large, and its use so pervasive, that in an institution with as many students with LD as Boston University there certainly must be several students who received their LD label by what amounts to a statistical mistake. Many critics of the LD field have commented that this statistical mistake results in disproporionate accommodations and resources for individuals of high socioeconomic status (see Note 2).

**Violation of the Specificity Assumption**

It is widely accepted that a modular deficit can co-occur with high- or low-functioning centralized processes (the type of processes that constitute general intelligence; see Note 3). As discussed previously, the concept of LD (and its associated assumption of
specificity) literally requires such a modular organization of cognitive skills in order to guarantee conceptual coherence for the LD idea. Nevertheless, the specificity assumption—the conceptual foundation of the entire LD field—is often violated in theory and practice. For example, Siegel’s description of the BU case files (in this issue) lists a depressing litany of nonmodular processes. Likewise, Kelman and Lester (1997) described several conceptually incoherent (because nonmodular) disabilities recognized by the Stanford University Disability Resource Center, such as “memory deficit” and “reasoning deficit.” The latter is defined as “(a) trouble thinking in an orderly logical way, (b) difficulty prioritizing and sequencing tasks, and (c) difficulty applying learned skills to a new task” (p. 165).

This disability strikingly calls to mind a thought experiment I proposed several years ago (Stanovich, 1993) in which I hypothesized a disability termed dysrationalia. I argued that the disability seemed literally forced on the field by its persistent appeal to discrepancy-based definitions. Dysrationalia was defined as

the inability to think and behave rationally despite adequate intelligence. It is a general term that refers to a heterogeneous group of disorders manifested by significant difficulties in belief formation, in the assessment of belief consistency, and/or in the determination of action to achieve one’s goals. Although dysrationalia may occur concomitantly with other handicapping conditions (e.g., sensory impairment), dysrationalia is not the result of those conditions. The key diagnostic criterion for dysrationalia is a level of rationality, as demonstrated in thinking and behavior, that is significantly below the level of the individual’s intellectual capacity (as determined by an individually administered IQ test). (p. 503)

As I argued in the 1993 article, dysrationalia as a disability is more likely to meet the specificity assumption than the “memory deficit” or “reasoning deficit” defined by Stanford. If we recognize the latter, the former seems a logical necessity. The logic of discrepancy-based classification based on IQ test performance has created a clear precedent whereby we are almost obligated to create a new disability category when an important skill domain is found to be somewhat dissociated from intelligence. The debate about the status of LD in the domain of social skills provides a clear example (Hammill, 1990; Hazel & Schumaker, 1988). One logical corollary of this practice is that the less comprehensive IQ tests are, the more such domains there will be; conversely, the more comprehensive and exhaustive the set of skills tapped by IQ tests, the fewer candidates for additional discrepancy-based disability categories there will be. However, the former is more likely because almost all critics of IQ tests make the argument that these instruments ignore many important domains of cognitive-behavioral functioning.

If these critics are right—if IQ tests do leave a lot of things out—the number of discrepancy-based disabilities would be extremely large. This raises the question of why intelligence should be the benchmark when it encompasses so little. To put it another way, if, as Kelman and Lester (1997) argued, intelligence is considered so encompassing that it is virtually taken as the essence of personhood, the number of discrepancies from it should be small, and LD will not encompass a large class of people. If, in contrast, learning disabilities are widespread (as advocates are prone to argue), intelligence must indeed reflect a thin slice of cognitive functioning (discrepancies from it in a variety of domains being quite common), and we seem justified in asking what makes it the superordinate construct that trumps all other cognitive abilities. After all, it is reading that is deemed discrepant from intelligence, rather than the converse—we have been less concerned about those who read well but display low intelligence.

Unresolved Issues

The Sociopsychometric Politics of LD

In arguing against the university’s call for more recent assessments, the plaintiffs in the BU case were ignoring what I have termed the sociopsychometrics of learning disabilities. As amply documented by Kelman and Lester (1997) and in the commentaries of Siegel, Sparks and Javorsky, and Elswit, Geetter, and Goldberg in this issue, children throughout the United States are not routinely and uniformly screened for LD at a fixed time in their academic careers using standard techniques and instruments. Instead, the diagnosis of LD in the schools reflects a complex educational process thoroughly entwined with a host of sociopolitical agendas. It represents a delicate negotiation among the social, political, educational, and resource concerns of parents, administrators, teachers, school and private psychologists, and special education personnel (mediated by legislation, advocacy, and lobbying of various types; see Kelman & Lester, 1997). Sparks and Javorsky (in this issue) provide a good example of the sociopsychometrics of LD classification at the university level when they note that “these findings also suggest that many students in this study were classified as LD based only on their problems with foreign language coursework at this university.”

In my view, Boston University and other universities are correct to resist being held hostage to processes of diagnosis that are as much sociopolitical as they are scientific (Carrier, 1986; Coles, 1987). If universities desire to exert control over the allocation of accommodations (accommodations that necessarily redirect resources and rewards away from other students) by requiring a more uniform scientific procedure for classification, then requiring recent assessments from personnel meeting a high criterion of expertise seems a valid way of bringing scientific rigor to the process and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
a valid way of removing the politics of social class and school resource allocation from the process. The random histories of school resource allocation decisions as they are played out in hundreds of school districts are something that universities should be trying to remove from their process of accommodation granting. Recent assessment is no panacea, but it is a step in the direction of more uniform criteria administered by qualified personnel who are aware of the latest research developments and not held hostage by the politics of the LD designation in school systems or by the pseudoscientific history of the field.

**Theories of Resource Allocation**

Kelman and Lester (1997) faulted LD advocates for failing to explicitly outline which theory of social justice drives their advocacy, stating that we must “recognize that there are a number of philosophical perspectives about how to allocate educational opportunities that need to be explored in some depth” (p. 9). For example, they discussed the following three philosophical rationales for providing additional resources (or accommodations) to certain groups:

1. prior resource deprivation,
2. relative prospective benefits of spending more on one child than another, and
3. the need to create greater equality of outcome.

It is clear that, with regard to traditional discrepancy-defined learning disabilities, Alternative 3 is a nonstarter. Within this framework, advocacy does not aid all poor performers—only those with aptitude–achievement discrepancies. Obviously, the net should be cast more widely if narrowing the total range of outcomes is the goal. Likewise, within this framework, Alternative 1 often turns out, after empirical investigation, to be a nonstarter. It is clear from reading the Saris decision in the Guckenberger et al. v. Boston Uni-

versity case and the commentaries in this issue that the plaintiffs had no history of resource deprivation relative to other students. Instead, it seems likely that their history had been one of receiving extra resources. The philosophical rationale of discrepancy-based LD advocacy clearly rests on Alternative 2, relative prospective benefits of spending more on certain individuals. This is where the previously discussed implicit assumptions about IQ serving as a proxy for educational “potential” come in. But these tacit assumptions run headlong into empirical difficulties; there is no hard evidence indicating that poor readers with an aptitude–achievement discrepancy respond differently to additional educational inputs and accommodations than poor readers without such discrepancies. I do not intend to review the evidence here, only to point out that the LD field has yet to begin a discussion of the social justice implications of extra resources and educational accommodations for certain students and not for others based on considerations of aptitude–achievement discrepancy. The first step in such a discussion would be to acknowledge and be very explicit about the zero-sum resource allocation that all educational institutions face.

**Zero-Sum Resource Allocation**

The plaintiffs in the BU case might well argue in response to the previous section that an argument for an educational accommodation does not amount to the claim of Alternative 2 above, and that this is really not their rationale at all—that there are no comparative implications for other students. They would be wrong, however. Take the case of providing extra time on examinations at the university level—an accommodation that is discussed in the BU litigation. Providing such time to students with LD functionally raises their performance on these exams and hence their grades. As Kelman and Lester (1997) argued, there is every reason to believe that peers without LD would also benefit from the extra time—particularly those who read almost as poorly but who, because of subpar IQs or social backgrounds, never triggered a sociopsychometric diagnosis of LD.

Grades are a resource, and they are allocated on a zero-sum basis like every other educational resource. They are a positional, relative resource; raise everyone’s grade a level and nothing would change. The law schools and medical schools would still want the top 10%, whatever the numerical or letter designation happened to be. As Kelman and Lester (1997) noted, “given our ongoing supposition that grades are either explicitly curved or that students understand their grades only as positional, relative goods, accommodationism also means that grades for non-LD students will decline, given any level of academic talent and studying” (p. 189). They thus term the extra-time accommodation for students with LD a “lump sum head tax on being non-LD” (p. 191). Their language here, although jarring, is useful. First of all, it forces us to think in the zero-sum terms that are really appropriate to this situation. A consideration of the zero-sum logic also makes it clearer that a specific argument in terms of distributive justice is called for—and has heretofore not been articulated in the LD literature.

Kelman and Lester’s (1997) language helps us to decentre our thinking about the issue (Stanovich, 1996a, 1999), dislodging unexamined assumptions. In essence, they have reframed the problem (Schick, 1997; Tversky & Kahneman, 1981) in a language that primes different automatic responses in us. This is often a useful exercise, particularly in zero-sum situations where it is sometimes hard to see that one person’s gain is another’s loss. For example, it is rarely noted by anyone except economists that a tax deduction for citizens having a certain characteristic is the equivalent of a penalty to those not sharing the characteristic. The term “tax deduction” seems to
most people an unequivocally "good" term, but it is not. It is a two-edged term—carrying a benefit for those who qualify and a penalty (in the form of an increased burden of the fixed sum that the nation must pay for its services) for those who do not. Consider the tax deduction for home mortgage interest. On the surface it seems like a benign and good thing. It seems less benign when we describe it as "the rent payer's penalty." Likewise, a lower tax rate for capital gains sounds less benign when counterphrased as "the wage earner's penalty." Even the most popular and benign of all deductions, the tax deduction for dependent children, takes on a slightly ominous and cruel connotation when counterphrased as "the tax premium paid by the childless."

Kelman and Lester (1997) have performed an equivalent reframing by highlighting the zero-sum consequences of accommodations for any group of students, no matter how defined. Calling something an accommodation for LD will generate automatic support. Calling it a grade point increase for students with LD will generate some hesitation. Calling it a relative grade penalty for those individuals who do not have a learning disability puts the burden of proof entirely on the other side (see Note 4). This shifting of the burden of proof is justified in my view because of the LD field's history of thoughtless (and psychometrically error-prone) application of discrepancy-based classification and its failure to articulate the theory of distributive justice that warrants extra resources for discrepancy-defined individuals but not for other individuals with learning problems virtually as great.

The Continuous Nature of Reading Disability

The multidimensional continuity of the distribution of reading ability in general and for all its related cognitive subskills has been demonstrated empirically (e.g., Shaywitz, Escobar, Shaywitz, Fletcher, & Makugh, 1992). Reading disability does not represent a discrete condition. Instead, below-average reading merges gradually and imperceptibly into poor reading, very poor reading, and, finally, severely impaired reading. There is no "magic" point where one either has or does not have a disability. Siegel and I have discussed in this issue two slightly different—but equally arbitrary—places to cut the continuous distribution. However, classification, accommodations, and delivery of services tend to be discrete; one either receives an accommodation or one does not. Accommodations are not continuously graded to reflect the continuous nature of the distribution. If the 15th percentile represents a cutoff for classification in a particular situation, someone at the 13th percentile gets the full accommodation and someone at the 18th percentile receives nothing. Yet these two individuals are psychometrically and cognitively indistinguishable. The LD literature contains virtually no discussion of the anomalies and unfairness that result from imposing an arbitrary cut in a continuous distribution. If the LD field does not begin such a discussion soon, the BU case makes it clear that lawyers will.

Toward a Rational Accommodation Policy

Embarrassed as I am to single out scholars from outside the LD field (especially when they are lawyers!), I simply must point to Kelman and Lester's book Jumping the Queue: An Inquiry into the Legal Treatment of Students with Learning Disabilities (1997), which I have cited heavily here, as containing some of the best analysis of the conceptual and sociopolitical complexities surrounding the issue of accommodations for university-level students with disabilities. Their book is simply full of good sense. For example, with regard to time accommodations on exams, they argue that educational authorities controlling a particular examination situation must decide whether speed is to be considered a genuine academic virtue in the particular context. If speed is not considered a genuine academic virtue, there is no reason for it to be a factor in test performance, and every student should be given an untimed test. This will no doubt be the judgment in most cases, thus eliminating the difficult issues raised by giving extra time to only some classes of student. In those (probably very few) cases where speed is judged to be a genuine academic virtue, no time accommodation should be given to anyone.

A further sampling of what I consider good sense in Kelman and Lester's book would include their elaborations of the following arguments:

Decisions about resource allocation inevitably involve clashes among individuals, each of whom would almost surely benefit ... from more expensive, personally tailored education. The question of whether students with learning disabilities would benefit more from resource infusions than other pupils is, essentially, an empirical question. Current federal law does not permit local districts to engage that empirical question head on. At times, it appears to "answer" it by rhetorical fiat. (pp. 6-7)

When diagnosis is imprecise, people may receive a diagnostic label for irrelevant or impermissible reasons. Moreover, it may be less justifiable to sharply differentiate the benefits children with and without a certain label are entitled to if criteria used to categorize a child one way or the other are less exacting. (p. 10)

If we accept systems, such as typical grading systems, in which rewards are differentiated, can we justify rejecting outcome inequality based on differential performance that is purportedly a result of learning disability while accepting the inequalities based on distinctions in performance with other roots? (p. 13)

It simply does not get us very far to recognize that some group of people with learning disabilities exists. ... It is plausible, for instance, that all bad readers have phonological difficulties, and
if this is the case, it is hardly apparent why we would then organize an entitlement scheme around the learning disabled rather than the phonologically impaired. (p. 35)

Frequently, real, inborn, and practically identifiable traits exist which are nonetheless “punished” and do not give rise to remedial obligations. Most significantly, those born with mild, non-disabling “deficits” in the sorts of intelligence that are most generally socially prized are not accommodated on tests, immunized from discipline, or entitled by federal law to compensatory education. (p. 36)

The views of “potential” that are being constructed here are not enormously rationally sustainable—they ultimately focus not so much on the actual capacity of the student to perform as they focus on the ways in which she resembles others who are high performers. Moreover, the moral claim—that those most “like” others capable of high performance are more entitled to both rewards and inputs than others who perform poorly but don’t resemble high performers as much—is highly suspect as well. (p. 155)

In a world of limited resources, it is not enough to say that children with learning disabilities “deserve” resources; their claims inevitably compete with claims that could be made by other “deserving” pupils who can be described in a variety of ways (poor achievers, socio-economically disadvantaged, gifted but understimulated, and so forth)… We have little sympathy, though, for the position most compatible with IDEA’s treatment of those with learning disabilities: a position that fetishizes the IQ test and demands that resources be applied in such a way as to ensure that educational performance correlates as closely with IQ as possible. (p. 156)

It is sad to note that Kelman and Lester’s (1997) book, written by two people from outside the LD field, represents one of the most nuanced and rigorous discussions of how the psychometrics of ability differences interacts with the concept of LD and the sociopolitics of school and society. Nowhere in the regular LD literature is there a treatment of these issues that is as scholarly and as philosophically and scientifically sophisticated. LD advocacy will always be open to charges of “queue jumping” as long as the field refuses to rid itself of its IQ fetishism, refuses to jettison aptitude-achievement discrepancy, and fails to free clinical practice from the pseudoscientific neurology that plagued the field in the 1970s. President Westling’s comments and those of other critics (e.g., Shalit, 1997) may have been ill-informed, but they represent a wake-up call to a field that has shunned self-criticism and only reluctantly embraced the conceptual change reflective of fields that are responsive to scientific advance. A more inclusive definition of disability—one that abandons discrepancy notions—would represent a first step toward acknowledging recent scientific findings on the nature of reading disability.

Finally, it is important to note that none of the LD field’s critics are denying that reading disability is a real problem, both for the individual and for society. Many of these critics—myself definitely among them—come from the political left and would advocate substantially increased educational resources generally, as well as increased resources for all individuals with handicaps. However, advocacy for increased resources does not preclude the necessity of a full discussion of the rational allocation of those resources. That, in turn, requires a conceptual analysis of the difficult issues surrounding the identification of cognitive disabilities. Kelman and Lester (1997) correctly argued that it also requires a full discussion of the different models of distributive justice that might be used to justify resource allocation and accommodation decisions.

ABOUT THE AUTHOR

Keith E. Stanovich, PhD, is professor of human development and applied psychology at the University of Toronto’s Ontario Institute for Studies in Education. His research interests are in the psychology of reading, reading disabilities, the cognitive consequences of literacy, critical thinking, and decision making. Address: Keith E. Stanovich, Department of Human Development and Applied Psychology, Ontario Institute for Studies in Education, University of Toronto, 252 Bloor Street West, Toronto, Ontario, Canada M5S 1V6.

AUTHOR’S NOTE

The preparation of this article was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to Keith E. Stanovich. The peace and quiet of Port Quin and Doyden Point made the completion of this commentary possible.

NOTES

1. If I really controlled the world, I would require that each student’s assessment specialist be given a special version of one of my “author recognition tests” (Stanovich & Cunningham, 1993; Stanovich & West, 1989; Stanovich, West, & Harrison, 1995). My test would have five names on it: Linda Siegel, Joe Torgesen, Charles Perfetti, Sally Shaywitz, and Dick Olson. If examiners did not know any of these five names, I would not accept any of their assessments because I would not consider them competent due to lack of current knowledge of the rapidly developing reading disability field.

2. Kelman and Lester (1997, pp. 89–90) reported that the median family income of students receiving accommodated administrations of the SAT test was 50% higher than the median family income of all test takers—a clear instance of the socially advantaged leveraging further social advantages for their children. Minority students were 50% less likely to receive an accommodated administration of the SAT.

3. Reading disability is not defined by the presence of moderate to high intelligence—the mistaken inference constantly made by people touting discrepancy definitions. Instead, reading disability can occur throughout the intelligence continuum, as Siegel emphasizes in her commentary (in this issue).

4. I would reiterate that the penalty is particularly egregious for individuals who do not have a learning disability but who have failed to acquire the LD designation because of adverse socioeconomic backgrounds, lack of family influence, or low IQ.


