Discrepancy definitions of reading disability: Has intelligence led us astray?

Current definitions of reading disability or dyslexia all involve the existence of a discrepancy between reading ability and measured intelligence. It is argued here that the use of intelligence as an aptitude benchmark in the definition of dyslexia conceals illogical assumptions about the concept of potential. The author suggests instead the use of a more educationally relevant aptitude measure, such as listening comprehension. However, all discrepancy definitions predicated on mismatches between aptitude and achievement are called into question by findings that the acquisition of literacy fosters the very cognitive skills that are assessed on aptitude measures. These findings undermine the logic of discrepancy measurement by weakening the distinction between aptitude and achievement. The author concludes that the validity of a severe discrepancy between aptitude and achievement as the defining feature of dyslexia has yet to be established to a degree that would justify differential educational classification or treatment.

Des écarts entre aptitudes cognitives et habiletés en lecture dans la définition de la dyslexie: Aurions-nous fait fausse route?

Les définitions actuelles des troubles de lecture ou dyslexie postulent généralement des écarts entre le niveau d'aptitudes intellectuelles tel que mesuré par les tests standardisés et le niveau d'habiletés en lecture. Le point de vue défendu dans cet article est que le fait d'utiliser l'intelligence comme critère dans la définition de la dyslexie repose sur des postulats non fondés quant au concept de potentiel intellectuel. L'auteur propose d'utiliser plutôt des critères académiques, tels que la compréhension verbale. Malgré tout, toutes hypothèses prédissant des écarts entre des types d'aptitudes et des habiletés interreliées sont sujettes à caution dans la mesure où des données de recherche montrent que le développement des habiletés écrites influence le développement des aptitudes cognitives qui sont mesurées par les tests standardisés. Ces données jettent un doute sur la logique qui sous-tend ces hypothèses quant à l'existence d'écarts entre aptitudes cognitives et habiletés en lecture en minimisant la distinction entre aptitudes et habiletés. L'auteur conclut que l'existence de différences importantes entre aptitudes cognitives et habiletés en lecture comme critère de définition de la dyslexie reste à démontrer si l'on veut justifier le bien fondé de classements et de traitements différents des enfants dyslexiques.
Las definiciones de discrepancia en la dificultad de lectura: ¿Nos ha llevado la inteligencia por el camino equivocado?

LAS DEFINICIONES actuales de dificultades de lectura o dislexia envuelven todas la existencia de una discrepancia entre habilidad de lectura e inteligencia medida. Se argumenta aquí que el uso de la inteligencia como una marca de aptitud en la definición de dislexia esconde razonamientos lógicos acerca del concepto de potencial. El autor sugiere en cambio, el uso de una medida más relevante educativamente de medición de aptitudes, tal como la comprensión oral. Sin embargo, todas las definiciones de discrepancia predicadas en las uniones equivocadas entre aptitud y logro son cuestionadas por los hallazgos de que la adquisición de la lectura promueve las mismas habilidades cognitivas que se miden en las medidas de aptitud. Estos hallazgos minar la lógica de la medida de discrepancia al debilitar la distinción entre aptitud y logro. El autor concluye que la validez de una severa discrepancia entre aptitud y logro como la marca que define a la dislexia aún está por establecerse a un grado tal que justificara la clasificación educativa o el tratamiento diferencial.

Diskrepanzdefinitionen der Lesestörung: Hat uns die Intelligenz irregenehrt?

IN ALLEN derzeitigen Definitionen über Lesestörung oder Legasthenie ist eine Diskrepanz zwischen Lesefähigkeit und gemessener Intelligenz vorhanden. An dieser Stelle wird argumentiert, daß die Verwendung des Faktors Intelligenz—als eine Eignungsbestimmung bei der Definition von Legasthenie—die unlogische Annahme eines Potentialkonzepts verschleiert. Stattdessen schlägt der Verfasser vor, eine Eignungsmaßnahme anzuwenden, die stärker auf die Schulbildung anwendbar ist, wie z.B. das Hörverständnis. Alle Diskrepanzdefinitionen, die auf der Fehlanpassung zwischen Eignung und Leistung basieren, würden durch die Feststellung in Frage gestellt, daß der Erwerb der Lesefähigkeit genau diejenigen kognitiven Fertigkeiten fördert, die anhand von Eignungsmaßnahmen festgelegt werden. Der Verfasser schließt daraus, daß die Gültigkeit einer wesentlichen Diskrepanz zwischen Eignung und Leistung als das definierende Merkmal für Legasthenie noch in gewissem Grade festgelegt werden muß, damit eine differenzielle Bildungsklassifizierung oder Bildungsmaßnahme berechtigt wäre.

The concept of dyslexia and/or reading disabil-ity has been controversial in the reading community. In this essay, I will argue that a major source of contention and theoretical confusion surrounding the term dyslexia stems from an almost perverse insistence on utilizing the concept of intelligence in definitions of reading disability.

As initially formulated, both professional and legal definitions of reading disability emphasized the existence of discrepancies between dyslexic children's actual school achievement and their presumed intellectual capacity. During the 1960s and 1970s, several proposed definitions of reading disability had considerable in-
fluence both on research issues and on debates about how best to deliver services to the children who needed them. For example, the World Federation of Neurology defined specific developmental dyslexia as "a disorder manifested by difficulty in learning to read despite conventional instruction, adequate intelligence, and socio-cultural opportunity. It is dependent upon fundamental cognitive disabilities which are frequently of constitutional origin" (Critchley, 1970, p. 11). Several features of this definition became canonical for many researchers and practitioners. But its "exclusionary criteria"—in particular, the requirement that a person have "adequate" intelligence to qualify for the dys-
lexia label—subsequently caused much dispute (e.g., Applebee, 1971; Ceci, 1986; Doehring, 1978; Eisenberg, 1978; Rutter, 1978).

The exclusionary criteria were carried over into the definition of learning disability employed in the landmark Education for All Handicapped Children Act (PL 94-142) passed in the U.S. in 1975:

Specific learning disability means a disorder in one or more of the basic psychological processes involved in understanding or in using language spoken or written, which may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations. The term includes such conditions as perceptual handicaps, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. The term does not include children who have learning problems which are primarily the result of visual, hearing, or motor handicaps, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage.

In specifically excluding children with poor reading achievement due to mental retardation, this definition highlights the requirement of a mismatch between aptitude and achievement.

The National Joint Committee for Learning Disabilities responded to criticisms of the exclusionary criteria by proposing that:

[]These disorders are intrinsic to the individual and presumed to be due to central nervous dysfunction. Even though a learning disability may occur concomitantly with other handicapping conditions (e.g., sensory impairment, mental retardation, social and emotional disturbance) or environmental influences (e.g., cultural differences, or inappropriate instruction, psycholinguistic factors), it is not the direct result of those conditions or influences (Hammill, Leigh, McNutt, & Larsen, 1981).

This response emphasized that the mere presence of other impairments or of environmental deprivation should not exclude children from being categorized as learning-disabled (see also Kavanagh & Truss, 1988).

All of these professional and legal definitions highlight the same salient feature: that a dyslexic child has an unexpected disability in the domain of reading, one not predicted by his or her general intellectual competence or socioeducational opportunities. Practically, this definition has been translated into a statistical assessment of the difference between a child's objectively measured reading ability and his or her general intelligence (Frankenberger & Harper, 1987; Kavale, 1987; Kavale & Nye, 1981; Reynolds, 1985; Shepard, 1980). Typically, very little effort is expended in ascertaining whether adequate instruction has been provided or whether the child suffers from sociocultural disadvantage—in short, in ascertaining whether the disability is "intrinsic to the individual." So much conceptual confusion has surrounded the discrepancy criterion that researchers and theoreticians have been reluctant to even consider the potential complications of the other criteria. In short, despite repeated admonitions that the diagnosis of reading disability should be multidimensional (Johnson, 1988; McKinney, 1987; Senf, 1986; Tindal & Marston, 1986), in actual educational practice the key defining feature is a discrepancy between reading achievement and aptitude as measured by an individually administered intelligence test (Frankenberger & Harper, 1987).

Intelligence as potential

The popularity of this use of the concept of intelligence as a benchmark in the diagnosis of reading disability is puzzling. Surely one would be hard pressed to find a concept more controversial than intelligence in all of psychology! For decades it has been the subject of dispute, which shows no sign of abating. Current research on individual differences in intelligent functioning continues to produce exciting findings and interesting theories (Baron, 1985; Ceci, 1990; Ceci & Liker, 1986; Sternberg, 1985, 1988), but still no consensual view of intelligence as a concept (Sternberg & Detterman, 1986). Even though much progress has been made in both empirical and theoretical domains, quite fundamental disputes remain. For example, some investigators have recently emphasized more contextualized approaches to the
study of intelligence (Ceci, 1990; Ceci & Liker, 1986; Sternberg, 1985, 1988; Sternberg & Wagner, 1986), whereas others have been advocating more decontextualized, biological approaches (Vernon, 1987).

Despite the controversy surrounding the concept of intelligence in the cognitive, developmental, and psychometric literature, measured intelligence was adopted as a foundational construct for the definition of dyslexia. Oblivious to the ongoing debates, specialists in learning disabilities seem to have avoided the issue by adopting a variant of E.G. Boring's dictum and acting as if "intelligence is what The Psychological Corporation says it is!" Teachers, schools, professional organizations, and government agencies accepted the choice of IQ test performance as the baseline from which to measure achievement discrepancies, in the absence of much critical discussion or research evidence. Educators and researchers have never grappled seriously with the question of why the benchmark should have been IQ. It is thus not surprising that so many of the conceptual paradoxes that plague the concept of dyslexia derive from the concept of intelligence (Stanovich, 1986a, 1986b, 1988b).

Why did professionals assent so readily to the use of IQ test scores in the discrepancy definition of dyslexia? Undoubtedly there were many reasons, but probably one factor was the belief that IQ scores were valid measures of intellectual potential. Certainly an extreme form of this belief is reflected in how dyslexia is portrayed by many advocacy groups and by the media. The typical "media dyslexic" is almost always a very bright child who is deeply troubled in school because of a "glitch" (assumed to be biologically based—see Coles, 1978, 1987; McGill-Franzen, 1987) that prevents him or her from reading. The subtext of such a portrayal is that the tragedy of the situation is proportionally greater because the child's great potential remains unlocked. This media portrayal has now entered the realm of folk belief, for there exists a popular myth that dyslexia is the "affliction of geniuses" (Adelman & Adelman, 1987; Coles, 1987)—that it is, if anything, more likely to occur in very bright people. This popular belief in the idea of unlocked potential undoubtedly helped to fuel the rapid expansion of the learning disabilities field.

One major problem with this popular belief, however, is that most psychometricians, developmental psychologists, and educational psychologists long ago gave up the belief that IQ test scores measured potential in any valid sense. Indeed, standard tests in educational measurement and assessment routinely warn against interpreting IQ scores as measures of intellectual potential (Anastasi, 1988; Cronbach, 1984; Thorndike, 1963). At their best, IQ test scores are gross measures of current cognitive functioning (Detterman, 1982; Humphreys, 1979). Indeed, many theorists would dispute even this characterization. Siegel (1988, 1989), for example, attacks the representativeness of several typical IQ tasks, outlining the objections of many theorists. Without entering into the details of all of these debates, the point is that an IQ test score is not properly interpreted as a measure of a person's potential. Thus, to the extent that IQ scores were viewed as measures of potential, the practice of diagnosing dyslexia by measuring discrepancies from IQ scores was misconceived from the beginning. 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.

However, advocates of current practices might counter some of these points by arguing that a strictly empirical orientation would support current procedures despite conceptual difficulties. That is, an advocate of the status quo might argue that all of the philosophical and conceptual criticisms are beside the point, because measuring discrepancy from IQ in the current manner distinguishes a group of children who, cognitively and behaviorally, are sufficiently distinct that the use of current pro-
cedures is justified on empirical grounds. Here we are getting to the heart of many recent research disputes.

Construct validity of the concept of dyslexia

The vast majority of poor readers in the schools are, of course, not characterized by severe discrepancies between their reading ability and assessed intelligence (Eisenberg, 1979). Their below-average reading performance is predictable from their general cognitive abilities. They are what Gough and Tunmer (1986) term garden-variety poor readers, and they tend to be more numerous than discrepancy-defined poor readers. The critical assumption that has fueled theoretical interest in the dyslexia concept from the beginning, and that has justified differential educational classification and treatment, is that the degree of discrepancy from IQ is meaningful—that the dyslexic has reading difficulties that either (a) stem from problems different from those that characterize the poor reader without an IQ discrepancy, or (b) are so much more severe for the dyslexic that they constitute a qualitative difference.

In empirical studies, this assumption has been tested principally through two research designs. One such design is the reading-level match design, in which an older group of dyslexic children is matched on reading level with a younger group of nondyslexic children. Without entering into the methodological thicket surrounding the use of this design to infer causation (Goswami & Bryant, 1989; Jackson & Butterfield, 1989; Vellutino & Scanlon, 1989), I suggest that the procedure of profile comparison in a reading-level match design seems at least a minimally acceptable operational method for testing the cognitive differentiability of dyslexic poor readers. The logic here is fairly straightforward: If the reading subskills and cognitive characteristics of the two groups do not match, then they would appear to have arrived at their similar reading levels via different routes. In contrast, if the reading subskill profiles of the two groups are identical, then there would seem to be little reason to differentiate between dyslexic children and other poor readers in theory or in educational treatment. If dyslexic children are reading just like any other child who happens to be at their reading level, and are using the same cognitive skills to do so, they become much less interesting from a theoretical perspective.

The second major design used to test the assumption of a qualitative difference—one pertinent not only to theoretical issues but also to the educational politics of reading disability—is to compare dyslexic children with children of the same age who are reading at the same level, but who are not labeled as dyslexic because they have lower IQs. Adapting the terminology of Gough and Tunmer (1986), I have termed this approach the garden-variety control design (Stanovich, 1988a). Again, the inferences drawn are relatively straightforward: A finding that the reading subskill profiles of the two groups do not match would at least be consistent with the assumption that they are arriving at their similar reading levels via different routes. In contrast, a finding that the reading subskill profiles of the two groups are identical would certainly undermine the rationale for the differential educational treatment of dyslexic children and would again make dyslexic children considerably less interesting theoretically.

Unfortunately, only recently have substantial numbers of well-controlled studies employing the garden-variety control and reading-level match designs begun to appear. For years, the great majority of studies of dyslexia employed only chronological-age controls, a design that has low diagnosticity (Bryant & Goswami, 1986). It was not until the mid-1970s that we had data from the groundbreaking epidemiological comparison of dyslexic and garden-variety poor readers conducted by Rutter and Yule (1975), and only in the last 5 years have their data been supplemented by other garden-variety control investigations. Similarly, it is only recently that enough studies employing reading-level matches have accumulated to make the patterns in the data discernible.
Part of the reason for the malaise and soul-searching that periodically overtakes researchers and practitioners in learning disabilities (Coles, 1978, 1987; Lyon, 1987; Senf, 1986; Stanovich, 1989; Swanson, 1988; Vaughn & Bos, 1987; Vellutino, 1979) is that the field plunged ahead into the domains of educational practice and diagnosis without first setting itself on a firm foundation by unequivocally demonstrating the empirical differentiability that would establish validity for the construct of reading disability. We turn next to the evidence that does exist.

**Phonological-core variable-difference model**

However, before summarizing the current state of the evidence on the cognitive differentiability of dyslexic poor readers I will first outline a model that illustrates what our data should show, ideally, if current conceptions of dyslexia are correct. I have termed this research framework the *phonological-core variable-difference* model (Stanovich, 1988a).

The model rests on the assumption of specificity inherent in current definitions of dyslexia (see Hall & Humphreys, 1982; Stanovich, 1986a, 1986b). This assumption underlies all discussions of the concept of dyslexia, even if not explicitly stated. It is the idea that a child with this type of learning disability has a brain/cognitive deficit that is reasonably specific to the reading task. That is, the concept of dyslexia requires that the deficits displayed by such children not extend too far into other domains of cognitive functioning; otherwise, the constellation of abilities we call intelligence would also be impaired, reducing the reading/intelligence discrepancy that is central to all current definitions.

In short, the key deficit in dyslexia must be a domain-specific process (see Fodor, 1983), rather than a central cognitive mechanism with widely distributed effects. For this and other reasons, many investigators have located the proximal locus of dyslexia at the word recognition level (e.g., Bruck, 1988, 1990; Gough & Tunmer, 1986; Morrison, 1984, 1987; Perfetti, 1985; Siegel, 1985, 1988; Siegel & Faux, 1989; Stanovich, 1986b, 1988b) and have been searching for the locus of the flaw in the word recognition module.

Studies in the last 10 years have focused intensively on phonological processing abilities. It is now well established that children classified as dyslexic display deficits in various aspects of phonological processing: They have difficulty making explicit reports about sound segments at the phoneme level, they display naming difficulties, their utilization of phonological codes in short-term memory is inefficient, their categorical perception of certain phonemes may be other than normal, and they may have speech production difficulties (Ackerman, Dykman, & Gardner, 1990; Cossu, Shankweiler, Liberman, Katz, & Tola, 1988; Kamhi & Catts, 1989; Liberman & Shankweiler, 1985; Lieberman, Meskill, Chatillon, & Schupack, 1985; Mann, 1986; Pennington, 1986; Pratt & Brady, 1988; M.A. Reed, 1989; Snowling, 1987; Taylor, Lean, & Schwartz, 1989; Wagner & Torgesen, 1987; Werker & Tees, 1987; Williams, 1984, 1986; Wolf & Goodglass, 1986). Importantly, there is increasing evidence that the linkage from phonological processing ability to reading skill is a causal one (Bradley & Bryant, 1985; Bryant, Bradley, Maclean, & Crossland, 1989; Lundberg, Frost, & Peterson, 1988; Lundberg & Hoien, 1989; Maclean, Bryant, & Bradley, 1987; Perfetti, Beck, Bell, & Hughes, 1987; Stanovich, 1986b, 1988b; Treiman & Baron, 1983; Wagner, 1988; Wagner & Torgesen, 1987). Presumably, a lack of phonological sensitivity makes it very difficult for a child to learn grapheme-to-phoneme correspondences. Some evidence points to reciprocal causation as well (Ehri, 1987; Ehri, Wilce, & Taylor, 1987; Morais, Bertelson, Cary, & Alegria, 1986), but a full discussion of this issue is beyond the scope of this essay and is not necessary to the argument.

Overall, then, voluminous evidence has accumulated to indicate that phonological deficits are the basis of the dyslexic performance pattern. However, such a conclusion would be an oversimplification, because it would omit the possibility of core deficits in the realm of orthographic processing. In fact, there is growing evidence for the utility of distinguishing a group
of dyslexics who have severe problems in accessing the lexicon on a visual/orthographic basis (see Stanovich, in press; Stanovich & West, 1989). Suggestive evidence comes from work on acquired reading disability that has revealed the existence of surface dyslexia (Patterson, Marshall, & Coltheart, 1985), a condition in which the affected individual appears to have difficulty forming and/or accessing orthographic representations of words in memory. Indirect evidence comes from multivariate investigations that have indicated that efficient phonological processing is a necessary, but not sufficient, condition for attaining advanced levels of word recognition skill (Juel, Griffith, & Gough, 1986; Tunmer & Nesdale, 1985).

But two caveats are critical when drawing conclusions about visual/orthographic processing deficits in dyslexia. First, there is much evidence indicating that the number of children with such deficits must be much smaller than the number of children with phonological difficulties (Aaron, 1989a; Freebody & Byrne, 1988; Gough & Hillinger, 1980; Liberman, 1982; Liberman & Shankweiler, 1985; Pennington, 1986; Perfetti, 1985; Vellutino, 1979). For example, if the number of children with visual/orthographic deficits were large, then their performance profiles would have obscured the identification of phonological problems in samples that were not preselected for subtypes; this has not occurred. Second, it is important to distinguish the difficulties encountered by these children from the “visual perception” problems portrayed in the early history of the study of dyslexia—which are now widely recognized to have been overstated (Aman & Singh, 1983; Morrison, Giordani, & Nagy, 1977; Stanovich, 1986a; Vellutino, 1979). Actual problems with visual/orthographic processing experienced by some children must be much more subtle and localized than older views had suggested. Nevertheless, some research has indicated the presence of visual deficits of some kind (Lovegrove & Slaghuis, 1989; Martin & Lovegrove, 1988; Solman & May, 1990; Willows, in press), but the issue remains controversial (Hulme, 1988). Irrespective of how this research dispute is eventually resolved, any smaller group of dyslexics with deficits in the visual/orthographic domain would mirror the group with phonological-core deficits in all other cognitive characteristics.

In the framework I have called the phonological-core variable-difference model (Stanovich, 1988a), the term variable difference refers to the key performance contrasts between the garden-variety poor reader and the dyslexic poor reader. Research has indicated that the cognitive status of garden-variety poor readers appears to be described well by a developmental lag model; that is, their cognitive skills are remarkably similar to those of younger children reading at the same level (Stanovich, Nathan, & Vala-Rossi, 1986; Stanovich, Nathan, & Zolman, 1988). A logical corollary of this pattern is that the garden-variety reader will have a wide variety of cognitive deficits when compared to same-age controls who are reading at normal levels. It is important to understand that the garden-variety poor reader does share the phonological problems of the dyslexic reader, and these deficits appear to be a causal factor in their poor reading (Juel, 1988; Juel et al., 1986; Perfetti, 1985; Stanovich, 1986b). But for the garden-variety reader the deficits—relative to age-matched controls—extend into a variety of cognitive domains (see Ellis & Large, 1987), some of which (e.g., vocabulary, language comprehension) may also be causally linked to reading comprehension. Such a pattern should not characterize the dyslexic, who should have a deficit localized in the phonological core.

The phonological-core variable-difference model assumes multidimensional continuity for reading ability in general and for all its related cognitive subskills. That is, we can conceive of all of the relevant distributions of reading-related cognitive skills as being continuously arrayed in a multidimensional space and not distributed in clusters. There is considerable evidence from a variety of different sources supporting such an assumption of continuity (Ellis, 1985; Jorm, 1983; Olson, Kliegl, Davidson, & Foltz, 1985; Scarborough, 1984; Seidenberg, Bruck, Fornarolo, & Backman, 1985; Share, McGee, McKenzie, Williams, & Silva, 1987; Silva, McGee, & Williams, 1985; Vogler,
Baker, Decker, DeFries, & Huizinga, 1989). However, viewing the distribution of reading skills as a graded continuum does not necessarily render the concept of dyslexia scientifically useless. Rather, as Ellis (1985) has argued, one can draw an analogy between dyslexia and obesity: Everyone agrees that the latter condition is a very real health problem, despite the fact that it is operationally defined in a somewhat arbitrary way by choosing a criterion from a continuous distribution.

The phonological-core variable-difference framework thus portrays the differences between the dyslexic and the garden-variety poor reader as differences in degree, rather than differences in kind. Consider the following picture: As we move in multidimensional space from the dyslexic to the garden-variety poor reader, we move from a processing deficit localized in the phonological core to the global deficits of the developmentally lagging garden-variety poor reader. Thus, the actual cognitive deficits one finds will be variable, depending upon the type of poor reader who is the focus of the investigation. The deficits on one end of the continuum will consist of deficits located only in the phonological core (the dyslexic); the differences will increase in number as we run through the intermediate cases that are less and less likely to pass strict psychometric criteria for dyslexia. Eventually we will reach the part of the multidimensional space containing relatively “pure” garden-variety poor readers who will not qualify for the label of dyslexic (by either regression or exclusionary criteria) because they will have a host of cognitive deficits; these poor readers will have the cognitively immature profile of a developmentally lagging individual.

This framework provides an explanation not only for why almost all processing investigations of reading disability have uncovered phonological deficits, but also for why some investigations have found deficits in other areas (see Stanovich, 1988b). As viewed from the phonological-core variable-difference perspective, virtually all poor readers have a phonological deficit; additional processing deficits emerge as one drifts in the multidimensional space from “pure” dyslexics toward garden-variety poor readers. Presumably, the studies that find deficits extending beyond the phonological domain are those that use more lax psychometric criteria for sampling poor readers; as a result, they incorporate more children from the garden-variety area of the space.

The phonological-core variable-difference model provides a parsimonious but realistic way of conceptualizing dyslexia. The model preserves at least some of the clinical insights of practitioners within the learning disabilities field who prefer traditional assumptions, yet its emphasis on continuity removes some of the more objectionable features of the term dyslexia—features that have actually stigmatized the term within some reading research subcommunities. Parsimony aside, however, the current empirical evidence that bears on the model is, quite frankly, mixed.

Empirical evidence for cognitive differentiability

The data from investigations employing reading-level match designs was once a confusing mass of contradictions (see Stanovich et al., 1986), but has recently been considerably clarified. Olson, Wise, Conners, and Rack (1990; see also Rack, Snowling, & Olson, 1990) have recently completed a meta-analysis of these studies that explains some of the discrepancies in the literature. It appears that the cognitive profiles of discrepancy-defined dyslexic readers do not match those of younger reading-level controls. The dyslexics are actually inferior in the phonological processing domain (Aaron, 1989b; Baddeley, Ellis, Miles, & Lewis, 1982; Bradley & Bryant, 1978; Bruck, 1990; Holligan & Johnston, 1988; Kochner, Richardson, & DiBenedetto, 1983; Lundberg & Hoien, 1989; Olson et al., 1985; Olson, Wise, Conners, Rack, & Fulker, 1989; Siegel & Faux, 1989; Siegel & Ryan, 1988; Snowling, 1980, 1981; Snowling, Stockhouse, & Rack, 1986). Although there are some exceptions to this pattern in the literature (Baddeley, Logie, & Ellis, 1988; Beech & Harding, 1984; Bruck, 1988; Treiman & Hirsh-Pasek, 1985; Vellutino & Scanlon, 1987), most of these can be explained
by a variety of factors that Olson et al. (1990; see also Rack et al., 1990) discuss in their meta-analysis. The data from reading-level designs thus provide at least modest support for the construct validity of the concept of dyslexia.

The evidence from studies employing a garden-variety control design is less consistent. Although some of these investigations have supported the idea of qualitative difference (Aaron, 1987, 1989b; Ellis & Large, 1987; Horn & O’Donnell, 1984; Jorm, Share, Maclean, & Matthews, 1986; Rutter & Yule, 1975; Silva et al., 1983), others have demonstrated that it can often be difficult to differentiate discrepancy-defined dyslexics from garden-variety poor readers (Fredman & Stevenson, 1988; Siegel, 1988, 1989; Taylor, Satz, & Friel, 1979). Even Olson et al. (1989) failed to find a correlation within their sample of dyslexic twins between degree of discrepancy and the degree of phonological deficit; although the statistical test is not quite equivalent to a garden-variety control design, the finding is nonetheless troublesome.

Regardless of how one views the muddled research evidence from the garden-variety control designs, the very fact that the literature is so full of contradictions forces one conclusion: namely, that it is surprisingly difficult to demonstrate cognitive differences between poor readers of differing IQs. The difficulty is surprising because it is intelligence that is supposed to be the more encompassing construct. Consider, for example, some data recently published by Siegel (1988): Reading skill, rather than IQ, is the factor that best differentiates subject groups on such variables as visual processing, phonological processing, ITAP performance, cloze performance, sentence correction tasks, short-term memory tasks, working memory tasks, spelling, and even arithmetic performance. As a general cognitive probe, reading ability seems to be more a more sensitive indicator than IQ test performance!

Additional problems of construct validity

Such findings are cause for some soul-searching. Indeed, the empirical support for the concept of dyslexia is, if anything, even more incomplete than I have so far portrayed it, for there are still inadequate data to sustain several other foundational assumptions. For example, outside of the pioneering work of Lyon (1985), data are scarce on differential response to treatment. There are, for example, no good data indicating that discrepancy-defined dyslexics respond differently to various educational treatments than do garden-variety readers of the same age (Pressley & Levin, 1987; van der Wissel, 1987). This is not a trivial gap in our knowledge. Differential response to treatment is, in large part, the raison d’être of special education.

We are equally unenlightened on several other crucial issues. The data are contradictory on whether there is differential prognosis for reading. Rutter and Yule (1975) found differential growth curves for specifically disabled and garden-variety poor readers. The garden-variety poor readers displayed greater growth in reading but less growth in arithmetic ability than the specifically disabled children. However, this finding of differential reading growth rates has failed to replicate in some other studies (Bruck, 1988; Labuda & DeFries, 1989; McKinney, 1987; Share et al., 1987; van der Wissel & Zegers, 1985).

Until we have convincing data on such issues as differential response to treatment, the utility of the concept of dyslexia will continue to be challenged because the reading disabilities field will have no rebuttal to assertions that it is more educationally and clinically relevant to define reading disability without reference to IQ discrepancy (Seidenberg, Bruck, Fornarolo, & Backman, 1986; Siegel, 1988, 1989). For example, Share, McGee, and Silva (1989) argue that

[i]t may be timely to formulate a concept of reading disability which is independent of any consideration of IQ. Unless it can be shown to have some predictive value for the nature of treatment or treatment outcome, considerations of IQ should be discarded in discussions of reading difficulties. (p. 100)

In short, the learning disabilities field is still in need of large-scale experimental demonstra-
tions that, compared with garden-variety poor readers, discrepancy-defined poor readers respond differently to treatment and have a different prognosis. The field is also in need of further evidence that the reading-related cognitive profiles of these two groups are reliably different.

**Deeper conceptual problems: The issue of intelligence**

We are thus right back to the issue of why IQ scores should have been the benchmark used to measure discrepancy in the first place. Indeed, it is surprising that for so long the concept of intelligence was so little discussed in the literature on learning disabilities. Researchers and practitioners in the field seem not to have realized that it was a foundational concept for the very idea of dyslexia. Without a clear conception of the construct of intelligence, the notion of a reading disability, as currently defined, dissolves into incoherence.

But problems with the IQ concept are numerous. 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 in a comparison study will result in the identification of somewhat different subgroups of children as discrepant and will also alter the types of processing deficits that the children will display (Bowers, Stoffy, & Tate, 1988; Fletcher et al., 1989; Lindgren, De Renzi, & Richman, 1985; J.C. Reed, 1970; Shankweiler, Crain, Brady, & Macaruso, in press; Siegel & Heaven, 1986; Stanley, Smith, & Powys, 1982; Torgesen, 1985; Vellutino, 1978). Yet researchers' recommendations for the type of IQ score that should be used are all over the map.

For example, a very common recommendation that one finds in the research literature is that performance and/or nonverbal IQ tests be used to assess discrepancy (e.g., Beech & Harding, 1984; Perfetti, 1985, p. 180; 1986; Siegel & Heaven, 1986; Stanovich, 1986a; Thomson, 1982), because verbally loaded measures are allegedly unfair to dyslexic children. In complete contrast, in a recent issue of *Learning Disabilities Research* devoted to the issue of measuring severe discrepancy, Hessler (1987) argues 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)]

Still other researchers suggest covering all the bases by using full-scale IQ scores. This unprincipled amalgamation of the above diametrically opposed philosophies is sometimes recommended precisely because the field is so confused and so far from consensus on this issue (Harris & Sipan, 1985, p. 145). Finally, there is a sort of "either" strategy that is invoked by investigators who require only that performance or verbal IQ exceed 90 in dyslexic samples (e.g., Olson et al., 1985). As Torgesen (1986) has pointed out, the naturally occurring multidimensional continuum of abilities guarantees that such a criterion ends up identifying more discrepancies from performance IQ than from verbal IQ.

Are there any implications in all of these differing practices beyond just the confusion created? Indeed, there are. First, there are implied value judgments in the measure used as a proxy for aptitude. Second, there are implications for precisely the research domain reviewed earlier: the empirical differentiation of dyslexic children from poor readers who do not show a discrepancy between aptitude and achievement. I shall take up each of these implications in turn.

**Value judgments inherent in indices of potential**

It is rarely noted that the use of certain types of intelligence tests in the identification of dyslexia often conceals conceptions of "potential" that are questionable, if not downright il-
logical. Consider again some of the hidden assumptions behind the often-heard complaint that verbally loaded intelligence tests are unfair to dyslexic children, and the admonition that performance IQ measures provide “fairer” measures of the potential of such children. A typical argument is that “[t]he 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). A similar argument is that “[c]omputing 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 person’s potential to comprehend verbal material. On the contrary, verbally loaded measures seem likely to provide the best estimates of how much a dyslexic reader could get from written text if his or her deficient decoding skills were remediated. As Hessler (1987) notes,

[...] there are different types of intelligence, and they predict academic achievement differently. In fact, the performance score accounts for so little academic achievement that there is reason to question its relevance for use as an ability measure to predict academic achievement. It is therefore a mistake to use any test of intelligence as an ability measure for predicting academic achievement in a severe discrepancy analysis simply because it is called a test of intelligence, without demonstrating some ability to predict academic achievement. (p. 45; see also Lyon, 1987, pp. 78-79)

Consistent with this interpretation is van der Wissel’s (1987) finding that the extent to which an IQ subtest separates dyslexic children from children who are garden-variety poor readers is inversely related to how highly the subtest correlates with reading achievement (see, however, Elliott, 1989). It is a paradoxical situation indeed when the indicators that seemingly best delineate dyslexia are those least related to the criterion performance that drew professional attention to dyslexia in the first place: reading failure.

It goes largely unnoticed that, when people argue for the greater fairness of nonverbal tests, they in effect jettison the notion of potential, at least in its common meaning. They cannot mean the potential for verbal comprehension through print if the dyslexic’s decoding deficits were remediated, because this is not what IQ tests—particularly the performance tests they are recommending—assess. Instead, people who make the “fairness” argument are implicitly asserting that dyslexics could perform much better if society were not organized so much around literacy. That assertion is undoubtedly true, but it is based on a counterfactual premise and on an interpretation of potential that contradicts more common usages. It simply makes little sense to adopt a linguistic usage of the term potential that requires 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 test—as a circumscribed behavioral index without loading it with social, and indeed metaphysical (Scheffler, 1985), baggage. If IQ tests are mere predictors of school performance, then let’s treat them as such. If we do, then performance IQ is manifestly not the predictor that we want to use, at least for predicting reading performance. An alternative conception of potential to apply in cases like dyslexia—in which educational achievement in a particular domain is thwarted due to a circumscribed, modular dysfunction—has already been alluded to: the degree of improvement that would result in a particular educational domain if the person’s dysfunction were totally remediated.

Implications of the choice of IQ measures for differentiating dyslexic from other poor readers

The aptitude measure used to differentiate dyslexic from garden-variety poor readers also has strong implications for the likelihood of isolating a modular dysfunction—perhaps in the phonological domain (Liberman & Shankweiler, 1985; Stanovich, 1988b)—that
could conceivably have some neurological localizability and genetic specificity. The question is this: Do we really want to look for a group of poor readers who are qualitatively differentiable in terms of etiology and neurophysiology? Officials at the National Institutes of Health (NIH) in the U.S. who are funding several program projects on the neurological, genetic, and behavioral underpinnings of dyslexia certainly do want to look for such a group (Gray & Kavanagh, 1985). Many who study learning disabilities share this enthusiasm for the quest to isolate—behaviorally, genetically, and physiologically—a select group of "different" poor readers.

Let us, for the purpose of the argument, accept this as a goal whether we believe in it or not. If we do aim to isolate such a group, we are led to the somewhat startling conclusion that we must move away from measures of abstract intelligence and toward more educationally relevant indices as benchmarks for establishing a discrepancy between aptitude and achievement. In short, to get the neurologically differentiable groups sought by NIH, we need an aptitude benchmark of more educational relevance than IQ—than nonverbal IQ, in particular, contrary to some common recommendations.

The context for any such discussion must be the voluminous body of prior research on the cognitive correlates of individual differences in reading achievement. The cognitive patterns that can be observed in studies of dyslexia are severely constrained by what we already know about the structure of human abilities in this domain. For example, a large body of research has demonstrated that reading skill is linked to an incredibly wide range of verbal abilities. Vocabulary, syntactic knowledge, metalinguistic awareness, verbal short-term memory, phonological awareness, speech production, inferential comprehension, semantic memory, and verbal fluency form only a partial list of the verbal abilities that have been linked to reading (Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Byrne, 1981; Carr, 1981; Chall, 1983; Cunningham, Stanovich, & Wilson, 1990; Curtis, 1980; Evans & Carr, 1985; Frederiksen, 1980; Harris & Sipay, 1985; Jackson & McClelland, 1979; Just & Carpenter, 1987; Kambh & Catts, 1989; Palmer, MacLeod, Hunt, & Davidson, 1985; Perfetti, 1985; Rapala & Brady, 1990; Rayner & Pollatsek, 1989; Siegel & Ryan, 1988, 1989; Stanovich, 1985, 1986a; Stanovich, Cunningham, & Feeman, 1984; Stanovich et al., 1988; Vellutino, 1979; Vellutino & Scanlon, 1987).

In contrast, the nonverbal abilities linked to reading are much more circumscribed (Aman & Singh, 1983; Carr, 1981; Daneman & Tardif, 1987; Hulme, 1988; Lovegrove & Slaghuis, 1989; Siegel & Ryan, 1989; Stanovich, 1986a; Vellutino, 1979; but see Carver, in press, for an opposing view). Moreover, the nonverbal abilities associated with reading are more likely to be distinct and domain-specific (e.g., orthographic storage, processing of certain spatial frequencies), whereas the verbal abilities related to reading are more likely to have global influence (e.g., inferential comprehension, verbal STM, vocabulary), thereby affecting general verbal IQ. Therefore, dyslexics who are matched with nondyslexics on performance IQ will be likely to have broad-based deficits in the verbal domain because verbal IQ is allowed to become unmatched. But even if there are visual/orthographic deficits linked to reading disability, the converse is not true. Because there are not as many reading-related nonverbal processes and because those that do exist will certainly be more circumscribed than something like vocabulary or verbal memory, verbal IQ matching will not necessarily result in a sample of dyslexic subjects with severely depressed performance IQs.

We will now picture traveling across the continuum of potential aptitude candidates for discrepancy measurement with this research context and the goal of differentiating dyslectic children in mind. It immediately becomes apparent that defining reading disability by discrepancies between reading achievement and performance IQ will make it extremely difficult to differentiate dyslectic children cognitively from other poor readers. Because they are allowed to have depressed verbal components, such performance-discrepancy dyslexics will have a host of verbal deficits, some of which
will occur at levels higher than phonology; consequently, these performance-discrepancy dyslexics will not display the cognitive specificity required of the dyslexia concept. Torgesen (1986) has discussed how the same effect will result from requiring only that either the verbal or the performance scale be above some criterion value. The verbal scale is allowed to be considerably below that of the nondisabled control group, so it is not surprising that a broad range of verbal deficits are subsequently observed (see Vellutino, 1979). A behaviorally and neurologically differentiable core deficit will be virtually impossible to find, given such classification. As Bowers et al. (1988) note,

ought the subtle language processing deficits differentiating dyslexics from normal readers have considerable overlap with verbal intelligence, then it will be impossible to distinguish the causes of dyslexia from the causes of poor reading associated with lower general verbal ability. (p. 307)

In contrast, discrepancies based on verbal aptitude measures would be likely to isolate a more circumscribed disability that may be more readily identifiable by neurophysiological and/or genetic methods. Such a procedure would preclude the possibility of identifying poor readers with deficits in broad-based verbal processes. It could potentially confine deficits exclusively to the phonological module. For example, Bowers et al. (1988) demonstrated that, if only performance IQ was controlled, dyslexic subjects differed from nondyslexics on rapid naming performance and on digit-span and sentence memory. Controlling for verbal IQ removed the association between reading disability and memory abilities, leaving only the association between reading disability and rapid naming performance. In short, verbal IQ control resulted in the isolation of a more circumscribed processing deficit.

Similarly, a discrepancy measure based on verbal IQ would be much more likely to demarcate a visual/orthographic deficit, if one exists (see Lovegrove & Slaghuis, 1989; Solman & May, 1990; Willows, in press). Matching on verbal IQ in a comparison study would, of course, allow the performance IQs of the dyslexic group to fall below those of the control group. But because the number of nonverbal abilities linked to reading is much more circumscribed than the number of such abilities in the verbal domain, the groups would not become unmatched on a large number of abilities. Moreover, subtle, visual-based deficits would not be “adjusted away” by matching on performance IQ. Thus, verbal IQ control provides a greater opportunity for comparison studies to locate such visual/orthographic deficits—which are much harder to track than deficits in the phonological domain.

In summary, by adopting verbal IQ as an aptitude measure we would be closer to a principled definition of potential in the reading domain as the academic level that would result from instruction if the person’s dysfunction were totally remediated. We would also be more likely to isolate a circumscribed deficit shared by a group of children who could be differentiated cognitively and neurologically from other poor readers, unlike the samples of children that have been defined by other methods.

Can we do without IQ? An alternative proposal

But if we have come this far down the road toward altering our measure of aptitude, then why not go all the way? The seeming fixation on intelligence in defining learning disability has driven a wedge between groups of investigators in the reading research community (McGill-Franzen, 1987). We have, for example, two organizations concerned with reading—the Orton Society and the International Reading Association—that hold conventions with almost totally nonoverlapping attendees. This strange state of affairs can be attributed directly to the marked differences between subgroups within the reading community in their views of the concept of dyslexia.

There is, however, a proposal for a way of conceptualizing reading disability that could well result in a rapprochement between groups like IRA and the Orton Society. In fact, this proposal has been around for quite some time, but has never received a proper hearing because
studies and definitions of dyslexia have so strongly emphasized the measurement of intelligence. In fact, many educationally oriented reading researchers have long suggested that it would be more educationally relevant—and would have been a more logical choice in the first place—to measure the discrepancy between reading ability and listening comprehension (see Aaron, 1989a; Carroll, 1977; Carver, 1981; Durrell, 1969; Gillet & Temple, 1986; Gough & Tunmer, 1986; Hood & Dubert, 1983; Royer, Kulhavy, Lee, & Peterson, 1986; Spring & French, 1990; Sticht & James, 1984). Calculating a discrepancy from listening comprehension certainly would seem to have more face validity and more educational relevance than the traditional procedure (Aaron, 1989a; Durrell & Hayes, 1969; Hoover & Gough, 1990; Spache, 1981). After all, children who understand written material less well than they would understand the same material if it were read to them appear to be in need of educational intervention. Presumably, their listening comprehension exceeds their reading comprehension because their word recognition processes are inefficient, causing a “bottleneck” that impedes comprehension (Gough & Tunmer, 1986; Perfetti, 1985; Perfetti & Lesgold, 1977). Reading comprehension correlates with listening comprehension much more highly than it correlates with full-scale IQ or even verbal IQ. Thus, children who are simultaneously low in reading and listening ability do not have an "unexplained" reading problem (Carroll, 1977; Hoover & Gough, 1990)—and we must always remember that the idea of "unexplained" reading failure is the puzzle that enticed us into the idea of dyslexia in the first place.

Listening comprehension is superior even to verbal IQ in its usefulness for isolating a modular deficit; in a comparison study, dyslexics matched to nondyslexics on listening comprehension would not show deficits in a host of reading-related verbal abilities. Moreover, the possibility that visual deficits play a role in dyslexia—a hypothesis that, although out of favor at the moment, is still the "layman's conception of dyslexia" (not totally without support in the literature: see Lovegrove & Slaghuis, 1989; Solman & May, 1990)—would receive a fairer test if the criterion for subject selection were a discrepancy between reading comprehension and listening comprehension. Thus, not only would we have a better chance at demarcating the modular phonological deficits that are of great interest in current work on dyslexia (Liberman & Shankweiler, 1985; Stanovich, 1988b), but we would also afford a fairer hearing to more tenuous hypotheses in the visual domain.

In short, in subjects who show a large reading discrepancy from listening comprehension we have probably isolated—as closely as we ever will—a modular decoding problem. Once isolated, this problem may or may not prove amenable to the kind of genetic and neurological analysis currently being pursued in NIH program projects. It seems ironic that measuring discrepancies from listening comprehension—a procedure often suggested by those hostile to the concept of dyslexia—may be just the procedure that allows those working from a neurological perspective to succeed in their quest.

There are, of course, several practical obstacles to the measuring of discrepancies between reading and listening comprehension. For example, although several measures of listening comprehension ability have been published (Carroll, 1972, 1977; CTB/McGraw-Hill, 1985; Durrell & Hayes, 1969; Spache, 1981), it is possible that no individual measure has been sufficiently standardized across the range of ages, or has attained the psychometric properties, to serve as an adequate benchmark from which to assess discrepancy (Johnson, 1988). Other complications also arise, such as how to compensate for hearing problems or unfamiliarity with standard English. However, many of these problems are no more severe for listening comprehension measures than they are for certain IQ tests. It is encouraging that work on the measurement of listening comprehension as a diagnostic tool has been on the increase recently, and important progress is being made (see Aaron, 1989a; Carlisle, 1989; Hoover & Gough, 1990; Royer, Sinatra, & Schumer, 1990; Spring & French, 1990).
Some of the other complications that arise with use of this alternative measure in fact illu-
minate the arguments outlined earlier. For ex-
ample, what type of material should we use to
assess listening comprehension? One possibility
is to use text material and simply present it
orally (Hoover & Gough, 1990). Regular text
material is often used in oral presentations in re-
search investigations and is employed in most
standardized measures of listening. Such mate-
rial will have the syntax, vocabulary, and lan-
guage structures of text, which of course differ
somewhat from the corresponding aspects of
speech (Chafe & Danielewicz, 1987; Hayes,
would be to use more naturalistic, non-textual
language in the oral presentation. Which
method is preferable obviously depends upon
the question being asked. But if our goal is ulti-
mately to identify a qualitative difference that
characterizes reading-disabled children, then
the choice is orally presented text. Performance
on such material will correlate more highly with
reading comprehension; most importantly, any
discrepancies will result more specifically from
decoding problems. In contrast, a severe dis-
crepancy between reading comprehension and
naturalistic language comprehension could
result from more global difficulties with text
structures, including difficulties that result from
environmental conditions (e.g., lack of ex-
posure to print).

The issues underlying the choice between
naturalistic speech and textual material for mea-
suring listening comprehension are analogous
in form to the issues underlying the choice be-
tween IQ measures discussed earlier. We want
to use the measure that is more highly corre-
lated with reading, because such a measure will
more cleanly isolate an unexpected, separable,
modular failure in a more circumscribed cog-
nitive domain. The assumption that such a
domain-specific deficit can occur has been the
critical intuition that has fueled interest in dys-
lexia from the time of Hinshelwood (1895,
1917) to the present.

It is important to note that educational prac-
titioners may also want to demarcate children
who have high nonverbal abilities coupled with
low ability in both listening and reading, in
order to give them special attention (this is a
policy issue); however, such children should
not be considered dyslexic. They do not have a
domain-specific difficulty in the area of read-
ing if their general listening skills are also de-
pressed. They may well present an important
educational problem worth identifying and con-
fronting, but it is simply perverse to call them
reading-disabled. In a contentious field, one
of the few areas of agreement is that “the current
sine qua non of learning disabilities is unex-
pectedly low achievement” [italics added]
(McLeskey, 1989, p. 435). We surely do not ex-
pect children who do not comprehend spoken
language to read well. Children who cannot un-
derstand material that is spoken to them are not
likely to understand it when they read it (Car-
their poor reading ability is not unexpected,
such children are not reading-disabled. But
— and this cannot be emphasized enough — this is
not to deny that these children have an educa-
tional problem; rather, it is simply to call for
more logical consistency in our application of
educational terminology and in our classifica-
tion of children.

Additional complications: Matthew effects

There remains, however, a further obstacle
to measuring reading disability by reference to
a discrepancy between aptitude and achieve-
ment, no matter what indicator is used as a
benchmark for aptitude. Let us again consider
the argument against the use of verbally loaded
tests for discrepancy measurement. This argu-
ment stems from the assumption, whether tacit
or explicit, that the reading difficulties them-
selves may lead to depressed performance on
verbally loaded measures. One reason a decline
in performance on such measures would be
likely is because of the “Matthew effects” asso-
ciated with reading — that is, the tendency of
reading itself to cause further development in
other, related cognitive abilities, such that “the
rich get richer and the poor get poorer” (see
Stanovich, 1986b; Walberg & Tsai, 1983).
However, to recognize the adverse effects of
reading difficulty on more general cognitive
abilities weakens the distinction between aptitude and achievement, perniciously undermining the whole notion of discrepancy measurement. This line of argument serves to remind us that, in their logic, current definitions of learning disabilities have implicitly given all of the causal power to IQ. That is, it is reading that is considered discrepant from IQ, rather than IQ that is discrepant from reading. Such a picture is a vast oversimplification, because there are potent effects running in both directions.

Much evidence has now accumulated to indicate that reading itself is a moderately powerful determinant of vocabulary growth, verbal intelligence, and general comprehension ability (Hayes, 1988; Hayes & Ahrens, 1988; Juel, 1988; Share & Silva, 1987; Share et al., 1989; Stanovich, 1986b; Stanovich & West, 1989; van den Bos, 1989). These Matthew effects (reciprocal causation effects involving reading and other cognitive skills) highlight a further problem with discrepancy-based classification: Do we really want to withhold certain types of educational treatment from children whose poor reading is accompanied by equally substandard IQ (or listening comprehension), when we know that the poor reading may be at least one direct cause of the low IQ and/or deficient listening comprehension ability? Because of the existence of such Matthew effects, we must allow the possibility that poor listening comprehension or verbal intelligence could be enhanced by better reading.

Thus, Matthew effects are interrelated in some very complicated ways with the conceptual logic of discrepancy-based definitions of disability. It appears that any discrepancy-based conceptualization is therefore going to require considerable refinement, based on how Matthew effects alter the course of development. Thus, any conceptually justified classification based on a discrepancy—even from listening comprehension—will be maddeningly tricky to carry out in a principled fashion.

Conclusion: Is discrepancy measurement worth the effort?

Much of the best recent research in the learning disabilities field has a "back-to-basics" feel to it (see Ceci, 1986; Torgesen & Wong, 1986; Vaughn & Bos, 1987), because researchers have begun to recognize that the field somehow got ahead of itself—that educational practice simply "took off" before certain fundamental assumptions had been thoroughly investigated. Thus, researchers have had to double back to retrace crucial empirical and theoretical steps that were skipped during the field's rapid expansion.

The concept of dyslexia has had a confused cart-before-the-horse history, in part because too many practitioners and researchers accepted at face value claims that IQ tests were measures of special “unlocked potential” in particular groups of children with low reading achievement. We have seen that in the area of reading disability the notion of unlocked potential was misconceived, because it was defined in a way that did not relate to the critical prediction in this domain: the prediction of how much growth in reading comprehension ability would be expected if the decoding deficit that was the proximal cause of the disability were to be totally remediated.

An alternative for measuring aptitude/achievement discrepancies, the use of listening comprehension ability as the aptitude benchmark, appears superior to IQ assessment. Nevertheless, complications stemming from the increasing difficulty of differentiating aptitude from achievement as a child gets older will plague all definitions based on the discrepancy notion. Problems such as these have led to Siegel’s (1988, 1989) suggestion that reading disability be defined solely on the basis of decoding deficits, without reference to discrepancies from aptitude measures. Whether or not her proposal is adopted, the learning disabilities field is simply going to have to face up to the following implications of current research findings:

1. Defining dyslexia by reference to discrepancies from IQ is an untenable procedure.
2. Much more basic psychometric work needs to be done in order to develop a principled method for measuring discrepancy from lis-
taining comprehension or some other verbal aptitude indicator.

3. If the field is unwilling to do the spade work necessary to carry out recommendation #2, or deems the potential benefit not worth the effort, then the only logical alternative is to adopt Siegel's proposal to define reading disability solely in terms of decoding deficiencies, without reference to aptitude discrepancy.

The field of learning disabilities has been so carried away with the face validity of the notion that discrepancies between achievement and aptitude should be important (given our standard psychometric assumptions and theories) that it has completely forgotten where the burden of proof lies. Because there is considerable evidence that reading ability is distributed on a continuum, that reading disability is not a discrete condition, and that the multivariate relations between reading ability and a variety of cognitive skills form a continuous space (Ellis, 1985; Jorm, 1983; Olson et al., 1985; Scarborough, 1984; Seidenberg et al., 1985; Silva et al., 1985; Vogler et al., 1989), the correct null hypothesis is that there are no qualitative differences between dyslexic and garden-variety poor readers in reading-related cognitive subskills. The burden of proof is on those claiming that there are such differences, because these assumed differences have been used to justify a new diagnostic category and associated educational practices.

In short, we are still in need of data indicating that the cognitive processing of dyslexic and garden-variety poor readers reading at the same level is reliably different, data indicating that these two groups of poor readers have a differential educational prognosis, and data indicating that they respond differently to certain educational treatments. These data, of course, should have been presented in the first place—that is, prior to the rapid expansion of discrepancy-based learning disability as a diagnostic and educational category. That we lack such basic data at this late date in the development of the learning disability field speaks volumes about the current state of educational research and practice in this area.

REFERENCES


FRANKENBERGER, W., & HARPER, J. (1987). States' criteria and procedures for identifying learning disabled chil-
Discrepancy definitions of reading disability

STANOVICH

25


Footnotes

Portions of this manuscript were presented at the Joint Conference on Learning Disabilities, sponsored by the International Academy for Research in Learning Disabilities and NICHD, University of Michigan, Ann Arbor, in June 1989, and at the New South Wales Department of School Education Conference on Learning Difficulties, Sydney University, Sydney, Australia, in June 1990. The author would like to thank Ron Carver and two anonymous reviewers for extensive comments on an earlier version of the manuscript. Requests for reprints should be sent to Keith E. Stanovich, Department of Psychology, Oakland University, Rochester, Michigan 48309-4401.

The terms reading disability, dyslexia, and specific reading retardation are used interchangeably in this essay. No connotations of differential etiology or cognitive pattern are implied by the use of different terms. Indeed, exploring the evidence on differential cognitive patterns is one of my goals here. Additionally, this essay is focused on school-based educational classification. These issues have somewhat different histories in the literatures of neuropsychology and clinical practice (Doris, 1986; Hynd & Hynd, 1984; Monaghan, 1980), which are beyond the scope of this essay.

Received November 20, 1989
Revision received August 8, 1990
Accepted September 1, 1990