CLASSIFICATION OF LEARNING DISABILITIES: AN EVIDENCE-BASED EVALUATION

Jack M. Fletcher, University of Texas; G. Reid Lyon, National Institutes of Health; Marcia Barnes, University of Toronto; Karla K. Stuebing, University of Texas; David J. Francis, University of Houston; Richard K. Olson, University of Colorado; Sally E. Shaywitz, Bennett A. Shaywitz, Yale University

INTRODUCTION

The purpose of this paper is to review research on the classification of learning disabilities (LD). We begin by briefly reviewing the nature of classification research. Then we discuss the evolution of definitions of LD, making explicit the classification hypotheses from which these definitions derive. An extensive review of the evidence for these hypotheses will be provided for the three components of classification implicit in the federal definition of LD: discrepancy, heterogeneity, and exclusion. We will show that classification hypotheses involving discrepancy and exclusion as embedded in federal (and state) policy have at best weak validity, often representing inaccurate and outdated assumptions about LD. There is evidence for heterogeneity of LD, but some reorganization of the types of LD identified in the federal definition may be necessary. Throughout the paper we identify alternative approaches to classification and identification, including weaknesses in any psychometric approach to the identification of LD. We suggest that classifications based on *inclusionary* definitions that specify attributes of different forms of LD are more desirable than current *exclusionary* definitions. Inclusionary definitions permit a focus on identification procedures that are intervention oriented as well as a focus on prevention, both of which are desirable and could contribute to improved results in remediating LD.

WHAT IS CLASSIFICATION?

Classification is the process of forming groups from a large set of entities based on their similarities and dissimilarities. It is not the same as identification, which is the process of assigning entities to an established classification. Valid classifications can be differentiated according to variables not used to form the groups. They are also reliable and have adequate coverage, i.e., permit identification of the majority of entities of interest. In classification research, groups are formed and evaluated for reliability, validity, and coverage. All classifications are hypotheses about the independent variables. Classification researchers evaluate the reliability, validity, and coverage of a hypothetical grouping of interest (Fletcher, Francis, Rourke, Shaywitz, & Shaywitz, 1993; Morris & Fletcher, 1988; Skinner, 1981).

Classification is fundamental to science and practice. It is virtually impossible to identify components of science or practice, regardless of the discipline and epistemological orientation, that do not involve classification. Although ubiquitous, classifications are often implicit and not explicitly identified. As part of science, however, all classifications are hypotheses that need to be empirically evaluated. Whenever a set of dependent variables is compared in relation to a set of independent variables (e.g., memory performance in children with and without LD), there is an explicit test of the hypotheses motivating the dependent variables (e.g., memory is weaker in LD), but also an implicit test of the independent variables (i.e., criteria for identifying children with and without LD) that derive from a hypothetical classification (Morris & Fletcher, 1988).

Even classifications that seem more straightforward, such as those used for defining children with and without traumatic brain injury, represent hypotheses at the level of the independent variables. To continue the memory performance example, if groups with and without traumatic brain injury differ in memory performance, evidence accumulates for the hypotheses that (a) memory is impaired in children with traumatic brain injury and (b) the criteria for defining traumatic brain injury are valid. The latter evidence would support the hypothetical classification of children along dimensions of brain injury (loss of consciousness, duration of coma, neuroimaging findings). Such evidence could be used to expand the classification towards hypothetical definitions of levels of severity (mild, moderate, severe); this classification and the criteria that lead to identification of children into severity groups could also be systematically evaluated along multiple dimensions: cognitive functions, prognosis, and response to intervention. The capacity of the classification to account for all children with traumatic brain injury (coverage) and to validly discriminate traumatic brain injury from other forms of brain injury (e.g., strokes, tumors) could also be evaluated. The keys are to recognize that there is a classification, to make it explicit, and to evaluate its reliability, validity, and coverage. When variation occurs in cognitive function, prognosis, or response to intervention among individuals with different levels of severity of traumatic brain injury, we can establish that the hypotheses leading to selection of these dependent variables were valid, but also that (a) the classification of injury severity has validity, and (b) the criteria used to operationalize the definitions of injury severity have validity (Fletcher et al., 1993).

In the area of LD, classification occurs at multiple levels: in identifying children as LD or typically achieving; as LD versus mentally deficient; within LD, as reading versus math impaired. Across classes of putative childhood conditions that produce underachievement, LD is identified as a particular type of “unexpected” low achievement and is distinguished from types where low achievement is expected due to emotional disturbance, social or cultural disadvantage, or inadequate instruction (Kavale & Forness, 2000). From a classification perspective, these levels of classification and the notion of LD as a form of low achievement that is unexpected represent hypotheses that should be evaluated.

That there are multiple underlying classifications of LD that are essentially hypotheses has not been consistently recognized. When the criteria for identifying LD began to evolve into policy in the 1960s, there was little research on which to base the underlying classifications and resultant definitions. This situation has gradually changed over the past 30 years, but the research that has emerged has had little impact on policy at the federal, state, and local levels. Indeed, the persistence of common assumptions about LD, its classification, and the perpetuation of resultant identification procedures are surprising given what has been learned about these disorders (Lyon et al., 2001). As we turn to research on the classification of LD, the question of how classifications should change as knowledge advances will emerge as a challenge to the field.

Definitions of Learning Disabilities: Implicit Classifications Made Explicit

The evolution of definitions of LD can be traced to the turn of the last century and is closely linked to concepts of organically based behavioral disorders (Doris, 1993; Rutter, 1982; Satz & Fletcher, 1980). The concept of LD arose from observations of children who were hyperactive and impulsive, but for whom the cause of the disorder was not obvious. As these problems often occurred in children for whom there were a history or some other suspicion of a brain injury, it was often presumed that the cause of these unexpected behavior disorders was constitutional in origin. Thus, these children were described with terms such as organic driveness syndrome, minimal brain injury, and then in the 1960s, minimal brain dysfunction. The latter label, stemming from a meeting convened by the federal government in 1962 (Clements, 1966), recognized that many children with these behavioral difficulties also had difficulty mastering academic skills with associated processing difficulties despite adequate intelligence and opportunities to learn.

In a subsequent meeting in 1966 convened by the U.S. Office of Education (USOE; 1968), the concept of LD, as proposed by Kirk (1962), was formally defined and considered as inclusive of minimal brain dysfunction and related disorders. The notion of minimal brain dysfunction as a disorder not attributable to mental deficiency, sensory disorders, emotional disturbance, or cultural or economic disturbance was retained. Etiological terms were dropped and replaced by educational descriptors, although the notions of unexpectedness and the implicit attribution to constitutional factors were retained. Parental and professional advocacy efforts led to the provision of special education services through the 1969 Learning Disabilities Act. The legislative language in the 1969 Act later appeared in the Education for All Handicapped Children Act of 1975 (Public Law 94-142) and is now currently reflected in the 1997 reauthorization of the Individuals with Disabilities Education Act (IDEA). All these legislative proceedings used the 1968 definition of LD:

> The term "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, 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 disabilities which are primarily the result of visual, hearing, or motor handicaps, or mental retardation, or emotional disturbance, or of environmental, cultural, or economic disadvantage. (USOE, 1968, p. 34)

After P.L. 94-142 was passed and federal funds became available, states were expected to identify children with LD. It quickly became apparent that states needed assistance with criteria for identification of LD, leading to publication of the Procedures for Evaluating Specific Learning Disabilities in the *Federal Register* (USOE, 1977). These procedures recommended that LD be defined as:

> a severe discrepancy between achievement and intellectual ability in one or more of the areas: (1) oral expression; (2) listening comprehension; (3) written expression; (4) basic reading skill; (5) reading comprehension; (6) mathematics calculation; or (7) mathematic reasoning. The child may not be identified as having a specific learning disability if the discrepancy between ability and achievement is primarily the result of: (1) a visual, hearing, or motor handicap; (2) mental retardation; (3) emotional disturbance, or (4) environmental, cultural, or economic disadvantage. (USOE, 1977, p. G1082)

Although states vary considerably in the IQ and achievement criteria used to designate a child as LD, discrepancy is used in either the definition and/or criteria by virtually all states, with the use of an IQ test to establish "aptitude" equally common (Frankenberger & Fronzaglio, 1991; Mercer, Jordan, Alsop, & Mercer, 1996). Discrepancy is the only inclusionary criterion; all other criteria are exclusionary and indicate simply what LD is not. Although there was little research at the time validating classifications of LD based on IQ discrepancy, researchers, practitioners, and the public commonly assume that IQ discrepancy is a marker for a specific type of LD that is unexpected and categorically distinct from other forms of underachievement (Kavale & Forness, 2000; Mercer et al., 1996; Stanovich, 1993). These beliefs reflect the common observation of unexpected underachievement in children who seem bright and capable.

The reification of IQ discrepancy in public policy is clearly apparent in the definition of LD in the 1992 and 1997 reauthorizations of IDEA, which continued the 1968 definition and added the following criteria from the 1977 recommendations to states:

> (a) A team may determine that a child has a specific learning disability if:
>
> (1) The child does not achieve commensurate with his or her age and ability levels in one or more of the areas listed in paragraph (a) (2) of this section, when provided with learning experiences appropriate for the child's age and ability levels; and
>
> (2) The team finds that a child has a severe discrepancy between achievement and intellectual ability in one or more of the following areas: (i) Oral expression; (ii) Listening comprehension; (iii) Written expression; (iv) Basic reading skill; (v) Reading comprehension; (vi) Mathematics calculation; or (vii) Mathematics reasoning. (U.S. Department of Education, 1999, p. 12457)

IQ discrepancy is clearly a prominent classification hypothesis. Other components of the federal definition also reflect classification hypotheses. Here we note the *heterogeneity* hypothesis, where LD is represented as seven different types of unexpected low achievement that may overlap. In addition, there is the *exclusion* hypothesis, which suggests that low achievement in LD is different from low achievement due to (a) mental deficiency and sensory disorders; (b) emotional disturbance; (c) social, economic, and cultural disadvantage; or (d) inadequate instruction. In the next sections, we review each of these classification hypotheses.

Discrepancy Hypothesis

The IQ-achievement discrepancy criterion is the most controversial and best-studied component of the federal definition of LD. From a classification perspective, it is a hypothesis that children with poor achievement below a level predicted by an IQ score (IQ discrepant) are different from children with poor achievement consistent with their IQ score (low achievement). IQ-discrepant children with LD have been proposed to differ from low achievers who are not IQ discrepant on several dimensions, including neurological integrity, cognitive characteristics, response to intervention, prognosis, gender, and the heritability of LD (Fletcher et al., 1998; Rutter, 1989; Siegel, 1992; Stanovich, 1991). There is an extensive body of research that can be used to evaluate this hypothesis. Although virtually all of the published studies involve reading disabilities (RD), we address LD in other domains later in this paper.

Isle of Wight Studies

The IQ-discrepancy classification hypothesis is not without support. The earliest empirical evidence validating IQ discrepancy came from the Isle of Wight studies in the early 1970s (Rutter & Yule, 1975). In this epidemiological study of RD, Rutter and Yule (1975) administered the Performance IQ Scale of the Wechsler Intelligence Scale for Children (WISC) and measures of reading. They defined two groups using a regression-adjusted definition: *specific reading retardation,* representing children with reading scores two standard errors below IQ, and *general reading backwardness,* representing children with reading scores that were deficient, but within two standard errors of IQ. In examining the distribution of residualized scores, they found an over-representation of children with general reading backwardness in the lower tail of the distribution of reading scores, representing a “hump.” They also found evidence suggesting that the two groups of poor readers could be differentiated, thus accepting the existence of a group of children with specific RD:

> Reading retardation is shown to differ significantly from reading backwardness in terms of sex ratio, neurological disorder, pattern of neurodevelopmental deficits and educational prognosis. It is concluded that the concept of specific reading retardation is valid. (p. 195)

Is There A Bimodal Distribution?

The Isle of Wight studies were widely accepted because they seemed to support the IQ-discrepancy hypothesis. Since that time, more critical evaluation of this support has become necessary. Although methodological factors involving inadequate ceilings on the reading measures have been cited (van der Wissell & Zegers, 1985), the critical issue centers around the interest of Rutter and Yule (1975) in the question of whether specific forms of RD could be distinguished from reading failure attributable to all other causes. Given this hypothesis, no exclusionary criteria were applied and approximately 36% of the children in the group defined as backwards readers had known or suspected evidence of a neurological disorder; many also had IQ scores in the ranges associated with mental deficiency. At the time, Rutter and Yule (1975) wrote that “it could be argued that the association with general reading backwardness was to be expected on the grounds of the below average intelligence of that group of children” (p. 189). It is well known that the distribution of IQ scores in a population is bimodal when individuals are included who have sustained injury to the central nervous system (Robinson, Zigler, & Gallagher, 2000). Not surprisingly, epidemiological studies in Australia (Jorm, Share, Matthews, & Matthews, 1986), New Zealand (Silva, McGee, & Williams, 1985), Great Britain (Rodgers, 1983; Stevenson, 1988), and the United States (Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992) that either excluded or had fewer children with brain injury have largely failed to replicate the Rutter and Yule (1975) finding of a bimodal distribution. This finding can be attributed to the prevalence of neurologically impaired children on the Isle of Wight, many with mental deficiency (Fletcher et al., 1998).

Can IQ-Discrepant and Low Achieving Poor Readers Be Differentiated?

Rutter (1989) observed that the critical test of the classification hypothesis does not depend on the presence of a bimodal distribution. Rather, the question is whether differences can be found that meaningfully differentiate IQ-discrepant and low achieving groups, which is a classification hypothesis. More recent studies of the validity of this hypothetical two-group classification, reviewed by Aaron (1997), Fletcher et al. (1993), Fletcher et al. (1998), Siegel (1992), and Stanovich (1991), have provided mixed evidence for the validity of the two-group classification. Many comparisons yielded null results, whereas others demonstrated small but statistically significant differences between the two groups.

When the studies are examined, they can be broken into domains involving prognosis, response to intervention, neurobiological factors, behavioral characteristics, achievement, and cognitive correlates. The bulk of the studies involve the behavioral, achievement, and cognitive domains, which are addressed in three meta-analyses summarized below. There is also research examining prognosis, response to intervention, and neurobiological factors. All six domains can be examined as evaluations of the validity of a two-group classification of poor readers based on presence or absence of IQ discrepancy.

Prognosis

Rutter and Yule (1975) reported that children who were backwards readers (i.e., low achieving) actually showed more rapid development of academic skills than children who were reading retarded (i.e., IQ discrepant). As the reading and spelling skills of the backwards readers were lower at baseline, and children were not randomly assigned to the two groups, the greater advances may reflect regression to the mean. Francis, Shaywitz, et al. (1996) examined this question using data from the Grade 9 follow-up of children in the epidemiological, population-based Connecticut Longitudinal Project. In this project, reading skills were assessed yearly beginning in Grade 1. The population is now being followed as adults.

Francis, Shaywitz, et al. (1996) composed three groups of children based on Grade 3 WISC–R full scale IQ and reading tests: not reading impaired (NRI), IQ discrepant using a 1.5 standard error regression-based criterion, and low achieving (not discrepant, but reading below the 25th percentile). Comparisons of the reading development of the three groups on the composite score from the Woodcock-Johnson Psycho-Educational Test Battery (Woodcock & Johnson, 1979) showed no differences between the two groups with RD in the rate of growth over time or the level of reading ability at any age despite the fact that about half the children in the IQ-discrepant group received special education services. As expected, both groups of poor readers differed significantly from the NRI group in growth rate and reading ability at all ages.

In Figure 1, these comparisons are carried through Grade 12. Again, there are clearly no differences in growth rates or level of reading ability at any age despite an 18-point difference in IQ between the two groups of poor readers. There was also no evidence that the poor readers narrowed the gap. More than 70% of those who read poorly in Grade 3 read poorly in Grade 12, showing that without intervention, LD in reading is a chronic, lifelong condition. These findings parallel those of Share, McGee, and Silva (1989), who reported results from another large longitudinal study in New Zealand. They found that IQ was not relating to reading achievement within age bands (7, 9, 11, 13 years) nor did IQ predict change over time. Share et al. (1989) concluded, “It might be timely to formulate a concept of reading disability that is independent of IQ. Unless it can be shown to have some predictive value for the nature of treatment or treatment outcomes, considerations of IQ should be discarded in discussions of reading difficulties” (p. 99).

Response to intervention

In turning to treatment, several studies examined outcomes in relationship to different indices of IQ or IQ discrepancy. Aaron (1997) reviewed earlier studies that sometimes included comparisons of groups defined as LD and low achieving, observing that both groups made little progress in their reading development, even

Figure 1.Growth in reading skills by children from 6–18 years of age (Grades 1–12) in the Connecticut Longitudinal Study based on the reading cluster of the Woodcock-Johnson Psycho-Educational Test Battery. The children were identified at 8 years of age (Grade 3) as not reading impaired (NRI), reading disabled according to a 1.5 standard error discrepancy between IQ and reading achievement (RDD), or low reading achievement with no discrepancy (25th percentile; low achieving). The figure shows that growth in

the two groups with reading disability is similar (the growth curves are indistinguishable); that neither catches up to the NRI group; and that the differences between the NRI group and the two groups with reading disability are apparent well before Grade 3.

with remedial placements. More recent studies explicitly examine this hypothesis in remedial or prevention efforts. In a remedial study of children with poor reading skills in Grades 2–5, Wise, Ring, and Olson (1999) assessed the relationship of full scale IQ in response to different approaches to intervention. They found that full scale IQ predicted about 5% of the variance in word reading outcomes on one measure of word reading, but that this effect was not apparent on other measures of word reading or assessments of phonological processing ability at the end of intervention. Similarly, Hatcher and Hulme (1999) found no relationships of IQ and reading outcomes involving word recognition.

Studies that have attempted to prevent RD in kindergarten and Grade 1 have also found no relationships of reading outcomes with full scale IQ or verbal IQ (Foorman, Francis, Beeler, Winikates, & Fletcher, 1997; Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Torgesen et al., 1999; Vellutino, Scanlon, & Lyon, 2000). Foorman et al. (Foorman, Francis, Beeler, et al. 1997; Foorman, Francis, Fletcher, et al. 1998) and Torgesen et al. (1999) examined relationships of reading intervention outcomes and general verbal ability, while Vellutino et al. (2000) looked both at levels of IQ and IQ discrepancy based on full scale IQ. In Vellutino et al. (2000), IQ-discrepancy scores were computed and compared among a variety of subgroups formed on the basis of reading gains, response intervention, and other indices. They concluded that "…the IQ-achievement discrepancy does not reliably distinguish between disabled and non-disabled readers … Neither does it distinguish between children who were found to be difficult to remediate and those who are readily remediated, prior to initiation of remediation, and it does not predict response to remediation" (p. 235). These findings are especially important in showing that IQ discrepancy is not specifically associated with those who respond to intervention.

In all the above studies, measures of phonological awareness skills were robust predictors of response to intervention. Some of these studies found that levels of IQ predicted growth in reading comprehension ability (Hatcher & Hulme, 1999; Torgesen et al., 1999; Wise et al., 1999), but consider what IQ tests actually assess. The subtests that make up a verbal IQ scale are commonly found to represent a general verbal comprehension skill closely related to vocabulary (Fletcher et al., 1996; Sattler, 1993; Share et al., 1989, 1991). As such, it is not surprising that IQ would predict reading comprehension as vocabulary is an essential part of IQ and a strong predictor of reading comprehension skills (Adams, 1990). Indeed, if IQ tests included measures of phonological awareness, it is likely that such measures would predict response to intervention. Inclusion of such subtests would also virtually eliminate the possibility that children with RD could ever be IQ discrepant given the close linking of phonological awareness skills and RD. Altogether, the results do not provide much support for differences in response to intervention between children defined as IQ-discrepant and low achieving poor readers.

Neurobiological factors

A series of studies from a group of researchers at the University of Colorado has been completed on the heritability of RD that addresses the validity of the IQ-discrepancy hypothesis. Pennington, Gilger, Olson, and DeFries (1992) classified a large population of monozygotic and dizygotic twins in which at least one member was classified with RD and a set of control twins in which neither was RD into one of four groups: RD based on IQ discrepancy, RD based on low achievement, RD based on both IQ discrepancy and low achievement, and those not classified as RD. Comparisons were made in three domains involving (a) genetic etiology, (b) gender ratios and clinical correlates, and (c) neuropsychological profiles. The researchers reported no evidence for differential genetic etiology based on type of definition. They also did not find evidence for significant differences in gender ratios, clinical correlates, and neuropsychological profiles.

More recent studies from this group have specifically tested the hypothesis that the genetic etiology of RD may vary by virtue of either IQ discrepancy or level of IQ. In a series of studies summarized by Wadsworth, Olson, Pennington, and DeFries (2000), genetic factors were more related to RD in children who have higher IQ scores than those with lower IQ scores. In Wadsworth et al. (2000), the overall heritability of reading disability was 0.58. Separating children defined as RD with full scale IQ scores above or below 100 resulted in heritability estimates of 0.43 for the lower IQ group and 0.72 for the higher IQ group, a statistically significant difference. These results indicate that environmental influences are particularly salient as a cause of reading difficulties in children with lower IQ scores.

These differences in heritability, while statistically and practically significant, are relatively small. Several earlier studies of the cohort with smaller samples yielded differences that did not reach statistical significance. Wadsworth et al. (2000) required almost 400 pairs of twins in order to detect the difference. It is not accurate to suggest that, because of these differences, classifications based on IQ discrepancy have value for components of LD other than the etiology of RD. As the researchers noted, the relatively high IQ of children with RD could be related to a more intractable genetically-based reading failure despite strong environmental support for IQ and for learning to read, whereas those children with RD who have relatively lower IQ scores may have more pervasive deficiencies in cognitive development and reading that reflect broader environmental disadvantages. For example, children in the lower IQ group in Wadsworth et al. (2000) had homes where there were fewer books and where mothers had fewer years of education. The researchers argued against excluding lower IQ children from intervention or remediation because they did not meet an IQ-discrepant definition, suggesting that the greater impact of environment influences on RD in this group suggests the need for emphasizing environmental intervention. Unfortunately, the traditional use of IQ and achievement criteria for LD in determining access to services has exactly the opposite effect.

There are also studies of children with RD that use functional imaging methods, such as functional magnetic resonance imaging (fMRI), which are reviewed in detail in the section on constitutional factors. While no study has a sample that is sufficiently large to actually compare IQ-discrepant and low achieving poor readers, it is noteworthy that no studies include only those children with IQ discrepancy. There is no evidence from these studies that children who meet IQ-discrepancy and low achieving definitions of RD have different neuroimaging profiles.

Meta-analyses of behavior, achievement, and cognitive ability domains

There are three meta-analyses that address the validity of IQ-discrepancy classifications for children with RD in the behavior, achievement, and cognitive ability domains and that constitute the bulk of studies of the IQ-discrepancy classification (Fuchs, Fuchs, Mathes, & Lipsey, 2000a; Hoskyn & Swanson, 2000; Stuebing et al., in press). The three studies were completely independent, but addressed slightly different questions. Fuchs et al. focused on the question of whether "the reading performance of underachieving children with and without the learning disabilities label is the same or different" (Fuchs, Fuchs, Mathes, Lipsey, & Eaton, 2000b, p. 2). To address this question, they identified and coded 76 studies that evaluated reading skills in children who were poor readers with and without the LD label. Fuchs et al. (2000a, b) reported a large effect size (0.76) showing poorer reading by groups with the label of LD in reading (presumably IQ-discrepant) relative to groups presumed to be poor readers without the LD label.

Hoskyn and Swanson (2000) coded 19 studies that met stringent IQ and achievement criteria. They focused specifically on studies where cognitive skills were compared in groups formed of those with higher IQ and poor reading achievement (IQ-discrepant) versus those with both lower IQ and poor reading achievement. They found negligible to small differences on several measures of reading and phonological processing (range = –0.02 –0.29), but larger differences on measures of vocabulary (0.55) and syntax (0.87). The groups were more similar than different, leading them to conclude that "… our synthesis concurs with several individual studies indicating that the discrepancy … is not an important predictor of cognitive differences between low achieving children and children with RD" (p. 117).

Stuebing et al. (in press) explicitly addressed the validity of the IQ-discrepancy classification hypothesis for RD in behavior, achievement, and cognitive domains. They reported on 46 studies that compared groups composed of poor readers who met explicit criteria for IQ discrepancy and low achievement. In the latter study, simply possessing the label of LD was not adequate, but some specification of the criteria used to designate children as IQ discrepant or low achieving was required. Fuchs et al. required the *label* of LD, with a presumption of IQ discrepancy, and some type of often unevaluated comparison group that presumably represented non-LD low achievers (e.g., placement in compensatory education). In contrast, Stuebing et al. required discrepancy criteria for the LD group and an indication that the low achieving group did not include individuals who might be IQ discrepant or typically achieving readers. These criteria were more liberal than Hoskyn and Swanson, but captured most of the 19 studies included in their meta-analysis.

Stuebing et al. (in press) found negligible aggregated effects for behavior (–0.05) and achievement (–0.12). A small effect size was found for cognitive ability (0.30). The effect sizes for the behavioral domain were homogeneous, but heterogeneity was apparent for the achievement and cognitive ability domains. When the heterogeneity was evaluated by examining the specific tasks within the achievement domain, those that involved word recognition, oral reading, and spelling showed small effect sizes indicating poorer performance by the IQ-discrepant groups. Tasks involving reading comprehension, math, and writing yielded negligible effect sizes. The small effect sizes for the former measures may reflect their similarity to the types of tasks used to measure poor reading in many studies. Similarly, constructs under cognitive ability closely related to reading yielded negligible effect sizes: phonological awareness (–0.13), rapid naming (–0.12), memory (0.10), and vocabulary (0.10). Not surprisingly, measures of IQ not used to define the groups yielded large effect size differences, while measures of cognitive skills like those measured by IQ tests (spatial cognition, concept formation) yielded small to medium effect sizes, the direction of both showing better performance by the IQ-discrepant group. Even with the inclusion of these measures of cognitive ability, the difference was only about three tenths of a standard deviation. Other analyses demonstrated (a) substantial overlap between the groups, and (b) that the size of the effects in different studies could be predicted by knowing the scores on the IQ and reading tasks used to define the groups (i.e., sampling variation across studies) and the correlation of these variables with the tasks used to compare the two groups. Stuebing et al. concluded that classifications of LD based on IQ discrepancy had at best weak validity.

The results of these three studies are quite consistent despite the differences in the research questions and the criteria for selecting studies. The most important difference was that unlike Stuebing et al. (in press), the other two meta-analyses did not differentiate IQ and achievement variables used to form the groups from those that served as dependent variables. It would be expected that variables used to define the groups would generate large effect sizes as IQ-discrepancy definitions select the poorest readers at each level of IQ (see Psychometric Issues below). To illustrate, Fuchs et al. (2000a, b) evaluated two constructs outside the reading domain that were not incorporated in the aggregated effect size estimate. The constructs yielded effect sizes consistent with Hoskyn and Swanson (2000) and Stuebing et al. (in press): 0.10 for phonological awareness and 0.26 for rapid naming. When measures of reading used to form groups were examined in Stuebing et al., a moderate effect size in reading showing poorer performance in children with IQ discrepancy was apparent. Altogether, these meta-analyses do not provide strong support for the validity of classifications based on IQ discrepancy.

Other Forms of LD and the IQ-Discrepancy Hypothesis

Discrepancy hypotheses have not received strong support in studies of RD, but LD is more than just RD. In this section, we review research on math disabilities, speech and language disorders, and psychometric issues relevant to any formulation of LD.

Specific math disability

As part of the Yale Center for Learning and Attention Disorders, Shaywitz (1996) evaluated the two-group classification hypothesis for computational disorders in math. The nature of these types of math disabilities (MD) is discussed below in the section on the heterogeneity hypothesis. Here we simply compare children who meet a 1.5 standard error IQ-discrepancy definition of MD with those who achieve below the 25th percentile, but whose math score on the Woodcock-Johnson Calculations subtest (Woodcock & Johnson, 1979) is within 1.5 standard errors of what would be predicted based on their full scale WISC–R score (Wechsler, 1974). These children do not meet criteria for RD using either IQ-discrepancy or low achieving criteria. They differ in full scale IQ (IQ-discrepancy $M = 107$, $SD = 12$; low achieving $M = 96$, $SD = 9$) and in math calculations (IQ-discrepant $M = 78$, $SD = 10$; low achieving $M = 85$, $SD = 4$). The nature and direction of the differences are exactly what would be expected given the properties of IQ-discrepancy definitions, where at each level of IQ the lowest performing children are identified into the IQ-discrepant group. Note also the reduction in the standard deviation relative to the population SD of 15, which is a product of subdividing a continuous distribution (Cohen, 1983).

Figure 2 shows a comparison of these two groups of children on a set of cognitive variables involving attention, language, problem solving, concept formation, and visual-motor skills. As Figure 2 shows, the IQ-discrepant group has higher performance levels on all variables. Note that neither group shows the severe impairment in phonological awareness associated with RD (see Figure 3 below). The group that is low achieving in math is noticeably poorer in vocabulary despite average reading skills. The critical issue, as for RD, is not that the groups differ; such differences in level of performance are expected because IQ tests are used to define the groups, and IQ is moderately to highly correlated with each of the measures (e.g., vocabulary) used to evaluate the children. Rather, the question is whether the pattern of differences separates the groups, implying that the correlates of math achievement differentiate the group. Testing the profiles for differences in shape did not yield a statistically significant difference and the effect size was negligible (0.06). As we have shown in the reading area (Fletcher et al., 1998), eliminating variability due to the difference in vocabulary eliminates the differences in level of performance apparent in Figure 2. The differences in Figure 2 are a product of the definitions and the correlates of poor math achievement do not appear to differ once the differences induced by the definition are taken into account.

Comorbid reading and math disability

Figure 3 compares IQ-discrepant and low achieving children with RD and MD on the same variables as Figure 2. In the upper panel, children with RD and no MD are depicted for contrast purposes, while the lower panel shows children with both RD and MD. In both panels, the striking impairment in phonological awareness is apparent. Note also the dip in vocabulary skills that characterizes both the low achieving groups. In the low achieving group that has only RD, the performance level in vocabulary is comparable with that of the low achieving MD group in Figure 2. Vocabulary is lowest in the low achieving RD-MD group. Again, these patterns reflect in part the relationship of IQ and vocabulary as opposed to specific associations with either RD or MD. The comorbid RD-MD group is more impaired in language skills, but also shows impairment on some of the same measures as the group that is only MD.

Speech and language disorders

Disorders of oral expression and listening comprehension are included under the LD category, though speech and language disorders are also a separate category in special education under IDEA. Epidemiological studies directed by Bruce Tomblin have explored the validity of IQ-discrepancy definitions in children who have disorders of expressive and receptive language. These comparisons have not supported the validity of IQ-discrepancy hypotheses for children with oral language disorders.

To illustrate, Tomblin and Zhang (1999) used measures of nonverbal IQ and oral language ability to create three groups of children from their large epidemiological study: not impaired, specific language impairment (IQ > 87 and composite language skills < 1.25 standard deviations below age), and general delay (IQ d” 87 and composite language skills < 1.25 standard deviations below age). Comparisons of the three groups on a variety of expressive and receptive language measures showed that the two language-impaired groups differed on multiple dimensions from the non-impaired group. Differences between the two language-impaired groups were less robust: “children with general delay closely parallel the specifically language-impaired group except that the children with general delay were more impaired and noticeably poorer on the test involving comprehension of sentences (grammatical understanding)” (p. 367). The investigators go on to question whether even this latter difference in grammatical understanding is specific to either group, noting, “current diagnostic methods and standards for specific language impairment do not result in a group of children whose profiles of language achievement are unique.” A consensus group convened by the National Institute of Deafness and Communication Disorders reached a similar conclusion (Tager-Flusberg & Cooper, 1999).

Psychometric Issues

Although we could continue a research program to evaluate the IQ-discrepancy hypothesis across multiple permutations, psychometric factors make it unlikely that any form of discrepancy can be effectively used. These factors raise questions about the viability of any approach to LD identification based solely on the use of test scores and cut-off points. Whereas to this point we have addressed the *validity* of LD classification, psychometric factors raise questions about the *reliability* of LD classifications.

Figure 4. Bivariate distribution of simulated IQ and achievement measures with a mean of 100, standard deviation of 1.5, and correlation of 0.6. Cutoffs depicting a 1.5 standard error discrepancy and low achievement (< 25th percentile) are drawn. Four segments are apparent: not reading impaired, only low achievement, only IQ discrepant, and both low achievement and IQ discrepancy.

In Figures 4–6, we examine what happens when groups are formed using IQ-discrepancy definitions in simulated data constructed to follow the bivariate normal distribution with no true group structure. It is apparent that the instability in these "simulated groups" parallels the instability seen in true groups (Shaywitz et al., 1992), raising doubts about the validity of the "true groups" formed by IQ-discrepancy rules. Consider Figure 4, which plots the bivariate distribution of simulated ability and achievement measures with a mean of 100, standard deviation of 15, and correlation of 0.6, consistent with population estimates (Sattler, 1993). Figure 4 also shows the groups that emerge when a 1.5 standard error regression definition like that employed in Connecticut (see Figure 1) is used, along with an arbitrary cutoff for low achieving at the 25th percentile. In Figure 4, it is clear that the groups are clearly demarcated, with no overlap in group membership. Note that many data points are below the low achieving cutoff, but are not IQ discrepant. Another subgroup is below both the IQ-discrepancy and low achieving cutoffs. A few children are above the low achieving cutoff but below the IQ-discrepancy cutoff.

Figure 5. Bivariate distribution of simulated IQ and achievement measures with a mean of 100, standard deviation of 1.5, and correlation of 0.6. Cutoffs depicting a 1 standard deviation discrepancy and low achievement (< 25th percentile) are drawn. The subject designations are from Figure 4 and show how simulated cases shift across the four segments by virtue of the change in the definition of discrepancy.

Figure 5 shows what happens when a different definition of discrepancy (one standard deviation) is used, analogous to how discrepancy is defined in many states, i.e., discrepancy without adjustment for the correlation of IQ and achievement. The symbols for the group represent their original locations in Figure 4. Note that the IQ-discrepancy cutoff is much steeper; the regression line in Figure 4 is actually slightly curved so that it is steeper at lower levels of IQ and flatter at higher levels of IQ. As a consequence, the unadjusted discrepancy definition identifies fewer children with lower IQs as discrepant (14% become low achieving) and identifies more children with higher IQs as "disabled" (6% of a large NRI group). The arbitrariness of the two discrepancy cutoffs is illustrated by asking what could possibly be the important differences in the 14% of children who change from IQ discrepant to low achieving at lower levels of IQ and other children who stay in these segments? Similarly, are the 6% of those who become "disabled" in Figure 5 truly impaired in reading? Fletcher et al. (1998) found no evidence supporting this hypothesis.

Figure 6. Simulated stability of group designations over time based on high stability (0.9) and reliability (0.8) for IQ and achievement measures. The subject designations are from Figure 4 and demonstrate the high instability associated with psychometric decision rules for identifying LD.

Figure 6 uses simulated data to show what happens to group membership over time. This figure was generated assuming high stability (0.9) in the traits measured and high reliability (0.8) for the measures of both traits. These assumptions mean that the traits vary little from person to person over time (i.e., individual differences are stable), and the traits are well-measured by the specific instruments. Thus, although there may be growth in the traits, growth does not differ much from one person to the other. These conditions should lead to a high degree of stability in classifications. Heterogeneity in growth would lead to instability in both individual differences over time and the classifications.

Figure 6 shows that classifications are not stable over time, despite the generally favorable conditions for stability. The instability is apparent in all four segments of Figure 6. In the group that is both IQ discrepant and low achieving, 38% move to the low achieving segment and another 38% move to the NRI segment. For the segment that is low achieving at Time 1, 14% move to the both IQ-discrepant and low achieving segment and 36% move to the NRI segment. In the Time 1 NRI segment, 3% move to the both IQ-discrepant and low achieving segment, 7% to the low achievement segment, and 1% to the IQ-discrepant-only segment. Finally, 67% of the only-discrepant segment moved to the NRI segment.

The lack of stability is also apparent when IQ and achievement scores are modeled from the Connecticut Longitudinal Study (Shaywitz et al., 1992). If IQ discrepancy and low achievement formed distinct and valid groupings, then one would expect stability in classifications over time, or at least instability that does not parallel the instability found in arbitrary classifications in simulated data. That IQ-discrepancy and low achieving classifications show instability that parallels the instability of arbitrary classifications in simulated data suggests that the IQ-discrepancy and low achieving distinctions are similarly arbitrary classifications, formed within a bivariate normal space and whose properties are largely driven by psychometric characteristics of this space rather than any inherent characteristics of the groups being formed.

Conclusions: Discrepancy Hypothesis

Concerns about the validity of the IQ-discrepancy classification hypothesis have led some to essentially reject the concept of LD (Ysseldyke, Algozzine, Shinn, & McGue, 1982), leading to fierce disagreements on whether LD and low achieving groups differ—all in defense of the concept of LD, not the validity of a hypothetical classification (Algozzine, Ysseldyke, & McGue, 1995; Kavale, 1995; Kavale, Fuchs, & Scruggs, 1994). The question is not so much whether children defined as IQ discrepant and low achieving are different, but how much they differ and whether the differences are meaningful for research and practice. The evidence reviewed above for prognosis, response to intervention, neurobiological factors, behavior, achievement, and cognitive abilities suggests that the IQ-discrepancy classification hypothesis lacks strong evidence for external validity. The psychometric evidence shows that the classification has problems with reliability. The criteria derived from the two-group classification produce groups of underachievers that are significantly overlapping; the differences that emerge are not strongly related to academic performance or to treatment and prognosis. Differences in behavior and cognitive abilities independent of the criteria used to form the groups are negligible. There is evidence for differences in the heritability of RD, but the differences are small, difficult to attribute solely to genetic factors, and with little evidence supporting the need to single out the IQ-discrepant group. There is no evidence from neuroimaging studies of a need to differentiate the groups; such studies routinely combine IQ-discrepant and low achieving children. Thus, consistent with the call of many researchers, the viability of the IQ-discrepancy classification hypothesis must be questioned.

Heterogeneity Hypothesis

In federal and non-federal definitions, LD is rarely conceptualized as a single disability, but instead is represented as a general category composed of disabilities in any one or a combination of several academic domains. In the 1968 federal definition, seven domains are specified: (1) listening; (2) speaking; (3) basic reading (decoding and word recognition); (4) reading comprehension; (5) arithmetic calculation; (6) mathematics reasoning; and (7) written expression. While the inclusion of these seven areas of disability in the federal classification ensures that the category of LD accounts for a wide range of learning difficulties, the practice implies that what may be highly variegated learning problems should be lumped together. Even today, many studies simply define groups of children as "learning disabled" despite considerable evidence that the correlates of LD in reading, math, and other achievement domains vary at multiple levels of analysis. In this section, we will ask how well these seven domains cover the range of LD and raise questions concerning what domains should be included in the federal definition.

Listening and Speaking

Disorders of listening and speaking are essentially oral language disorders. Such disabilities are incorporated in IDEA under the speech and language category, so the need for including them as types of LD category is not clear. As oral language disorders, they represent examples of difficulties with expressive and receptive language. What is the point of duplication, especially since disorders of listening and speaking are not formal areas of academic achievement? Difficulties in listening comprehension typically parallel problems with reading comprehension (Shankweiler et al., 1999; Stothard & Hulme, 1996). Children cannot understand written language any better than they can understand oral language. Any phonological, syntactic, or semantic problems that hinder oral language comprehension will also affect the ability to read written text or even to comprehend when someone reads them the text. While some children with LD have oral language disorders, the duplication is far from perfect (Tomblin & Zwang, 1999). The basis for including disorders of listening and speaking in the federal classification of LD is not clear and leads to conceptual confusion in classifying and defining oral language disorders in IDEA.

Reading Disabilities

The federal definition specifies two areas of reading difficulties, basic reading (word recognition) and reading comprehension. That difficulties with word recognition represent a specific form of LD in reading is well established (Shaywitz, 1996). Children can also be identified with comprehension difficulties that do not involve the word recognition module. Much more is known about the nature and causes of disabilities in word recognition, as less reading research has been devoted to studying how children understand what they read.

What are not addressed in the federal definition are difficulties that involve the automatization of word recognition skills and speed of reading connected text. These problems also occur in children with accurate word recognition skills. Unfortunately, less is known about fluency deficits in reading despite recent development of hypotheses suggesting that deficiencies in reading fluency represent a separate subgroup of RD (Wolf & Bowers, 1999; Wolf, Bowers, & Biddle, 2001). In the next section, we review evidence for subgroups with RD specific to word recognition, comprehension, and fluency.

Word recognition (dyslexia)

Word-level RD is synonymous with dyslexia, a form of LD that has been described during the 20th century as word blindness, visual agnosia for words, and specific reading disability (Doris, 1993). The evolution of the concept of dyslexia, and its link with word-level RD, provide an excellent example of how definitions of LD can move from exclusionary to inclusionary. As an example of an exclusionary definition, consider the 1968 World Federation of Neurology definition that was in part the basis for the epidemiological studies of Rutter and Yule (1975):

> A disorder manifested by difficulties in learning to read despite conventional instruction, adequate intelligence, and socio-economic opportunity. It is dependent upon fundamental cognitive disabilities, which are frequently of constitutional origin. (Critchley, 1970, p. 11)

In contrast, consider the following definition of dyslexia formulated by a research committee of the International Dyslexia Society (Lyon, 1995; Shaywitz, 1996), which we have modified to be consistent with advances in research:

> Dyslexia is one of several distinct learning disabilities. It is a specific language-based disorder characterized by difficulties in the development of accurate and fluent single word decoding skills, usually associated with insufficient phonological processing and rapid naming abilities. These difficulties in single word decoding are often unexpected in relation to age and other cognitive and academic abilities; they are not the result of generalized developmental disability or sensory impairment. Dyslexia is manifest by variable difficulty with different forms of language, often including, in addition to problems reading, a conspicuous problem with acquiring proficiency in writing and spelling. Reading comprehension problems are common, reflecting word decoding

and fluency problems.

This definition identifies dyslexia as a word-level RD proximally caused by phonological processing problems. It is inclusionary because it clearly specifies that a child is dyslexic who has (a) problems decoding single words in isolation, and (b) difficulties with phonological processing. These constructs are easily measured. The difficulty, of course, is specifying the level of impairment that would be of sufficient severity to constitute a disability. The definition is directly linked to intervention and it is now well established that treatments emphasizing the development of word recognition skills improve reading achievement in these children (National Reading Panel, 2000; Swanson, 1999). It reflects the developmental origins of dyslexia, so that prior to the expected onset of word recognition skills, interventions addressing the development of phonological processing skills should prevent word recognition difficulties. There is considerable research support for this expectation (National Reading Panel, 2000; Snow, Burns, & Griffin, 1998). The definition clearly permits the identification of children who are at risk for dyslexia and also permits identification of children who do not respond to preventative interventions and who may need different forms of remediation. No mention is made of discrepancy and IQ tests are not required for identification. It stipulates that dyslexia is differentiated from mental deficiency and sensory disorders, but criteria for these differentiations would be included in the identification of these disorders in an overall classification of low achievement. No distinctions or stipulations concerning cause or etiology are made, including constitutional factors, and exclusions are not identified.

The definition reflects a view of dyslexia that is different from those found in the media, where dyslexia is viewed as a rare, exotic disorder characterized by unusual perceptual characteristics (e.g., seeing words and letters backwards). Dyslexia as defined here is the most common form of LD and has its origins in the language system (Shaywitz, 1996; Vellutino, 1979). Lerner (1989) reported that 80% of all children served in special education programs have problems with reading, while Kavale and Reese (1992) found that 90% of children in Iowa with the LD label had reading difficulties. Most children who have reading problems have difficulty with word-level skills. It may not be the only problem that these children experience, but it is the problem that makes them poor readers. Most children served in special education programs as LD likely have word-level reading problems as part of their disability (Lyon, 1995).

Dyslexia as defined above is a disorder that is not associated with specific qualitative characteristics, but occurs on a continuum of normal development. Thus, dyslexia is the lower portion of this continuum (Shaywitz et al., 1992). A critical issue is where on the continuum sufficient severity of reading difficulty occurs that would lead to a designation of RD. This issue has not been adequately researched, but should be tied in some way to response to interventions of different kinds of intensity, not an arbitrary designation (e.g., 20th percentile) that the examples in Figures 4–6 show to be unreliable.

People with dyslexia often have other academic problems and also seem to have problems that are in the social and behavioral realm. This is not a problem with the definition, but with the classification of LD. The key is to have a classification that signals when a child has a form of LD, and which recognizes that they may have other academic and behavioral difficulties. Many children with this form of RD have problems with spelling, writing, reading comprehension, and math (Lyon, 1996). The spelling, writing, and reading comprehension problems can be explained on the basis of the disruption of phonological processing and word recognition skills. Spelling is closely tied to phonological processes; a person with poor word recognition skills cannot identify or spell words accurately because of poor understanding of the relationship of print and speech: the alphabetic principle. They will have reading comprehension problems because they can not process the text. When math is also impaired, the child typically has other problems involving oral language and working memory (Swanson & Siegel, in press). As we discuss below, reading comprehension and math problems in the absence of word recognition difficulties can also occur, which must be accounted for in our classification—not our definition of dyslexia.

Disorders like attention deficit hyperactivity disorder (ADHD) represent a different classification issue. While ADHD commonly co-occurs with dyslexia (Shaywitz, Fletcher, & Shaywitz, 1997), what is important is that the child with both dyslexia and ADHD looks dyslexic when their reading and language skills are examined and looks ADHD when their behavior is examined (Shaywitz et al., 1995). However, dyslexia is a problem with *cognitive* development; ADHD is a *behavioral* disorder with cognitive consequences (Barkley, 1997). Thus, the child has more than one disability, although children with both dyslexia and ADHD have more severe reading (and other cognitive) problems than children who have only dyslexia or ADHD. The treatment implication is that both disorders need to be addressed and that interventions addressing only one disorder may be less effective (Fletcher, Foorman, Shaywitz, & Shaywitz, 1999). ADHD is not a part of our classification of LD, as the primary defining characteristics do not reflect academic achievement.

Altogether, word-level RD is the best researched type of LD and the difference between the 1968 exclusionary definition and the modified 1994 inclusionary definition represent what we believe is a model for other forms of LD. As we see in the next sections, much progress needs to be made in other forms of LD, though we could formulate reasonable inclusionary definitions of most of these forms.

Reading comprehension disability

There is good evidence for disabilities in reading comprehension in cases where reading decoding is age-appropriate but reading comprehension lags. Estimates of the incidence range from 5% to 10% depending on the exclusionary criteria used to define the groups (e.g., Cornoldi, DeBeni, & Pazzaglia, 1996 vs. Stothard & Hulme, 1996). These estimates have not been studied in relation to age, but it is likely that specific reading comprehension problems are more apparent in older children and emerge after the initial stage of learning to read. Some may have a history of word recognition difficulties that have been remediated.

Studies on specific reading comprehension disability commonly have compared children with good word recognition accompanied by good reading comprehension skills with those who have good development of word recognition skills but poor development of reading comprehension (Nation & Snowling, 1998; Oakhill, Yuill, & Parkin, 1986; Stothard & Hulme, 1996). This is in contrast to studies that have investigated reading comprehension problems in groups that contain a large number of poor word decoders (e.g., Perfetti, 1985; Shankweiler et al., 1999), in which the sources of reading comprehension problems are difficult to address separately from the influences of difficulties in word decoding. Proficient reading comprehension presumes fluent decoding, so studies of reading comprehension must separately identify those weak in comprehension, but fluent in decoding.

IQ and the definition of reading comprehension disability. Research in the area of reading comprehension disabilities does not follow the classification guidelines that are embedded in the federal definition of LD. Most studies of children’s comprehension difficulties have not attempted to relate general intellectual ability to reading comprehension. Not surprisingly, there are few studies that use IQ-achievement discrepancies to define groups of poor comprehenders. The discrepancy formula that is most often used in studies of reading comprehension disability is that between good basic reading achievement and poorer scores on standardized tests of reading comprehension, without reference to IQ. Such approaches have not been fruitful, though most of the research is on children who also have word-level RD (Fletcher et al., 1998). One study that used an IQ-achievement discrepancy model to classify poor comprehenders found that children with average intelligence and average word reading skills but poor reading comprehension had difficulties in listening comprehension, in working memory, and in metacognitive aspects of comprehension (Cornoldi et al., 1996). A survey of individual cases showed that children with reading comprehension disability were heterogeneous with respect to the specific pattern of cognitive deficits that they displayed in these skills.

In some studies of reading comprehension disability, IQ has actually been used as an outcome measure rather than as an exclusionary criterion for group membership. For example, children with specific reading comprehension disability have been found to have similar phonological skills and nonverbal intelligence as children with no comprehension disability, but lower verbal IQs (e.g., Stothard & Hulme, 1996). Such findings have been interpreted by some as providing evidence that general verbal cognitive deficits underlie the reading comprehension disability of good decoders/poor comprehenders. In a recent study of normally developing readers, however, verbal intelligence was found to account for only modest variation in reading comprehension performance (Oakhill, Cain & Bryant, in press; also see Badian, 1999). After accounting for verbal intellectual skills, significant variance in comprehension was predicted by text integration skills, metacognitive monitoring, and working memory with stability in these relationships over a 1-year period. Interestingly, these are the same skills that Cornoldi et al. (1996) found best characterized their group of poor comprehenders with IQs that were discrepant from reading comprehension achievement.

What does it mean to say that children with comprehension problems have lower verbal IQ? A simple assumption of a unidirectional relationship between intelligence and comprehension, such that higher verbal intelligence somehow paves the way for the development of good reading comprehension, is probably incorrect for two reasons. First, there is some evidence that the relationship between reading comprehension and intelligence may be bidirectional (Francis, Fletcher et al., 1996). Consider, for example, that reading experience may facilitate growth of verbal and even nonverbal intellectual skills (Stanovich, 1993). Second, tests of verbal intelligence measure vocabulary and verbal reasoning, and these are some of the same skills that are measured by tests of reading comprehension. A moderately strong relationship between verbal intelligence and reading comprehension, then, is not unexpected and is relatively uninformative. Furthermore, given that there are important aspects of comprehension that IQ tests do not capture, verbal IQ cannot be used as a proxy for reading comprehension disability.

Core deficits in reading comprehension disability. Most of the research on specific reading comprehension disability has focused on determining the core deficits that underlie the disability. These studies have generally taken three forms. One is to compare children who are good decoders but poor comprehenders to good decoders–good comprehenders, matched for age. More recent studies use reading level match designs in which the cognitive processes of good decoders–poor comprehenders are compared to those of younger children matched for reading comprehension level to the older disabled children. Finally, studies of remediation have asked whether training in skills hypothesized to contribute to the reading comprehension deficit actually improves reading comprehension. The findings from the three methods are largely consistent and are summarized below.

Some studies have shown that children who are good decoders–poor comprehenders may have more basic deficits in vocabulary and understanding of syntax that would impair reading comprehension (Stothard & Hulme, 1992, 1996). Other studies have shown that even when vocabulary and syntax are not deficient, deficits in reading comprehension still arise (Cain, Oakhill, & Bryant, 2000; Nation & Snowling, 1998). The results from these studies are consistent with findings discussed previously (IQ-achievement discrepancy group in Cornoldi et al., 1996, and normally-developing readers in Oakhill, Cain, & Bryant, in press). These deficits involve inferencing and text integration, metacognitive skills related to comprehension, and working memory. In contrast, phonological skills, short-term memory, and verbatim recall of text are typically not deficient (reviewed in Oakhill, 1993; Cain & Oakhill, 1999; Cataldo & Cornoldi, 1998; Nation, Adams, Bowyer-Crane, & Snowling, 1999; Oakhill, 1993; but see Stothard & Hulme, 1992).

More recent studies in this area have begun to question how poor comprehension early in a child's reading history may influence not only later reading comprehension, but also continued development of word decoding skills. Although decoding and comprehension disabilities have been shown to be dissociable, children who are good decoders but poor comprehenders may begin to fall behind in their decoding skills in the later school grades (Oakhill, Cain, & Bryant, in press). To the extent that these individuals are not very good at using reading as a means to an end, they may come to read less, and so truncate their exposure to less common words (Cunningham & Stanovich, 1999). Alternatively, their poor ability to use semantic cues (a component of comprehension) to decode less frequent words may constrain higher levels of lexical development (Nation & Snowling, 1998).

Findings similar to those discussed for children with reading comprehension disability have also been found in studies of children with brain injury. For example, Barnes and Dennis (1992, 1996) have evaluated the discourse and reading comprehension skills of children with spina bifida and hydrocephalus. These children are often characterized by intact word recognition skills, but deficient reading comprehension (and math) abilities. Using a variety of tasks, Barnes and Dennis have demonstrated that children with this form of brain injury have difficulty making inferences and problems assimilating nonliteral information from text, and that these difficulties in the reading domain parallel problems that the children have in oral discourse comprehension and production. Table 1 summarizes the characteristics of reading ability in children with spina bifida and hydrocephalus, children with word recognition difficulties and poor comprehension, and non-brain injured children who have intact word recognition skills but poor reading comprehension.

Table 1. Academic subgroups of LD

1. Reading Disability—Word Level
2. Reading Disability—Comprehension
3. Reading Disability—Fluency (?)
4. Math Disability
5. Reading Disability and Math Disability
6. Written Expression—Spelling, text, handwriting (?)

The comprehension-related deficits outlined in Table 1 have been replicated across studies that have used different criteria for group membership, including brain injury. Questions remain regarding whether metacognitive, inferential, and working memory processes are primary causes or consequences of the comprehension deficit and whether difficulties in these skills reflect deficits in more basic reading comprehension processes (Nation & Snowling, 1998; Nation et al., 1999; Perfetti, Marron, & Folz, 1996). Given that reading comprehension may be a more multidetermined process than reading decoding, it is not unexpected that advances in knowledge in this area have lagged behind word-level RD.

Given that it is a multifaceted process, the assessment of reading comprehension is a major problem. In contrast to tests of word recognition accuracy in which there is a relatively transparent relationship between the content of the tests and performance requirements for word reading, there is more controversy about what reading comprehension tests measure. Standardized reading comprehension tests differ from everyday reading contexts along several potentially important dimensions such as passage length, immediate versus delayed recall, and learning and performance requirements (Sternberg, 1991). Reading comprehension tests, like other tests of complex cognitive functions, may be limited both by a lack of ecological validity and by the absence of a model of the reading comprehension process that would guide test construction. Thus, it is not surprising that there is less consensus on how to define reading comprehension disability and how to best advance understanding of the reading comprehension process in terms of both its normal and disordered development.

Reading fluency

More controversial is the question of whether there is a specific subgroup of reading impairment that is characterized specifically by difficulties in reading fluency. Wolf and Bowers (1999, Wolf et al., 2001) have argued for a “rate deficit” group that does not have problems in the phonological domain, but often has difficulties with comprehension because of a more general difficulty rapidly processing information. The subtyping study of Morris et al. (1998) did find evidence for a rate deficit subtype that was not phonologically impaired, but which showed difficulty on any task that required speeded processing, including rapid automatized naming and tasks as mundane as canceling target letters as fast as possible from an array of letters. This subtype also had difficulties with reading fluency and comprehension, but not word recognition.

Studies of ADHD show that reading fluency problems are common in these children with ADHD and that these difficulties are related to their performance on measures of rapid automatized naming (Tannock, Martinussen, & Frijters, 2000). Some argue that these difficulties reflect common underlying brain-based problems with timing or rapid processing that occur across all forms of reading disability, but more research needs to be completed (Waber, Wolff, Forbes, & Weiler, in press).

Studies of children with brain injury also provide evidence that the accuracy and speed of word recognition can and should be differentiated. Barnes, Dennis, and Wilkinson (1999) matched children with traumatic brain injury on their word decoding accuracy. Comparisons of reading rate and naming speed showed that fluency was worse in children with traumatic brain injury, paralleling observations with non-brain injured children with rate deficits (Waber et al., in press; Wolf et al., 2001). Fluency was related to reading comprehension scores in both populations (Barnes et al., 1999; Morris et al., 1998).

This discussion of rate deficits represents an excellent example of how classifications and definitions must evolve in supporting the provision of services for children with academic deficits. Although there may be insufficient evidence to establish a form of RD that involves only fluency deficits, the possibility is under active investigation. If evidence continues to accumulate for a fluency disorder, the classification must be changed and definitions of reading disability expanded to incorporate these types of problems.

It is also apparent that a single definition will not work for these three putative forms of reading disability. It is already possible to specify the attributes of each disability. These attributes can be measured at the level of the academic skill as well as its associated correlates. As we will discuss below, it may be possible to measure these attributes and form inclusionary definitions that lead to specific procedures for identification and have important implications for intervention.

Math Disabilities

The federal definition of LD specifies disorders of math calculations and reasoning. Disabilities in math have been studied for as long but not as extensively as RD. Nonetheless, there is a burgeoning research base, particularly on children who have computational difficulties. There is clear evidence for a specific subgroup of children with LD in math calculations; whether there is also a subgroup that has impairment in math concepts is unclear and hardly studied. It is even possible that there are other subgroups of MD that have yet to be determined. Consider that the skills that fall under the heading of mathematics are broad and varied and it is unclear whether learning in one domain of mathematics is related to learning in another domain (Geary, 1994). Unlike reading, in which development produces changes in quantity and quality of decoding and comprehension, the development of mathematical competencies involves learning new categories of skills such as geometry and calculus (LeFevre, 2000). These new skills depend to some extent on previously learned math knowledge, but these areas of math also represent significant departures from prior learning.

As in the reading area, studies of adults with brain lesions show that fairly specific math skills can be either preserved or lost depending on the damage to the brain (Dehaene, 1999). Whether the development of math skills across different domains can be similarly fractionated is an open question. In normal development, for example, the acquisition of basic arithmetic skills may facilitate the acquisition of more advanced math skills across a number of math domains (Geary, Fan, & Bow-Thomas, 1992). Because of insufficient knowledge about development of some areas of math, the cognitive skills that lead to competence in those areas, and the potential importance of computation in facilitating these other areas of mathematical development, this section will deal primarily with evidence for LD in basic arithmetic calculation.

Computational abilities and disabilities

There is a rich literature on the acquisition of skills such as counting, basic understanding of quantity, and use of strategies that are important to the development of early computation ability (e.g., Ashcraft, 1992; Bisanz, Morrison, & Dunn 1995; Gelman & Gallistel, 1978; Nunes & Bryant, 1996; Rourke, 1993; Siegler & Shrager, 1984). This work was not motivated by a need to understand disordered development of computational skills, but to understand cognitive development through mathematical cognition and to understand the development of the math system itself. In the past decade, some math researchers have used the theories and methods of mathematical cognition and developmental psychology to study the emergence and development of MD. Until recently, this work largely proceeded without respect to the specificity of the disability, that is, whether RD and MD were comorbid or specific (e.g., Geary, Bow-Thomas, & Yao, 1992; Geary, Brown, & Samaranayake, 1991; Jordan, Levine, & Huttenlocher, 1995).

In contrast, studies of children with LD, including those involving math computations, have indeed been concerned with describing differences between groups of children with either specific RD or MD and both RD and MD (e.g., Ackerman & Dykman, 1995; Fletcher, 1985; Morrison & Siegel, 1991; Rourke, 1993; Swanson & Siegel, in press; White, Moffitt, & Silva, 1992). Children with specific MD appear much less frequently than children who have both RD and MD, and precise estimates are not available (Rourke, 1993). The existence of such children is clearly established in many studies where children are defined as having word recognition difficulties, both word recognition and math computation difficulties, and only math computation difficulties. The latter children do not have problems with language of the sort experienced by children with word-level RD. They typically have difficulty with different forms of nonverbal processing and concept formation (Rourke, 1993). To summarize, these studies have found cognitive deficits in marker skills such as verbal and visual working memory and visual-spatial skill that differentially characterize the different subgroups of children with LD (see Figures 2 and 3). Although these studies reveal the importance of considering the specificity and comorbidity of learning disabilities, they do not permit an analysis of the mechanisms by which the cognitive marker skills influence math learning. Furthermore, such subtyping studies were interested in the issue of LD in math versus reading rather than in understanding the basic math processes that contribute to scores on math achievement tests (Ginsburg, Klein, & Starkey, 1998).

A recent development in the research on MD has been the attempt to combine research strategies from the fields of cognitive development, mathematical cognition, and LD (e.g., Geary, Hoard, & Hamson, 1999; Jordan & Hanich, 2000). These studies: (1) follow children longitudinally from the beginning of their school careers to understand how MD plays out over time in terms of issues such as stability and comorbidity; (2) measure the development of early, informal arithmetic skills such as counting and problem-solving strategies for computing numbers that may be related to the later development of school-based or formal mathematical competence (Ginsburg et al., 1998); (3) relate cognitive marker skills or cognitive competencies that are purported to support the development of math to the acquisition of specific components of the math system such as knowledge of counting principles (Geary et al., 1999); and (4) analyze components of the developing math system and their supporting cognitive competencies with respect to whether the MD is specific or comorbid with RD. In some ways this new research strategy parallels earlier longitudinal studies of reading acquisition in typically achieving children and those with reading problems. As these studies proved critical to our understanding of RD at multiple levels, the current studies may similarly begin to reveal the explanatory status of hypothesized core number-related skills and supporting cognitive competencies with respect to the classification and prediction of MD.

What are the core number-related skills and supporting cognitive skills that have been hypothesized to account for difficulties with mathematics calculations? Disability in math computation may arise from problems in learning, representing, and retrieving math facts from semantic memory and/or from difficulties in the acquisition and use of developmentally-mature problem-solving strategies or procedures to perform mental or written calculations (Geary, 1993). Whether difficulties in the spatial representation and manipulation of number information constitute a third source of MD in children with and without frank brain injury (e.g., Geary, 1993) is unclear and not well studied (Barnes et al., in press).

Comorbid reading and math disability

It has been suggested that the core deficit of children with both RD and MD might be difficulty in retrieving math facts from long-term or semantic memory (Geary, 1993). In its more general form (i.e., not just involving math fact retrieval, but also retrieval of lexical information), this memory-based deficit may also be related to some of the features of reading disability (Geary et al., 1999). Longitudinal research and treatment studies suggest that this type of computation deficit may not improve much with age or remediation (Geary et al., 1991; Goldman, Pellegrino, & Mertz, 1988). In terms of supporting cognitive skills, these types of math disabilities are hypothesized to relate to working memory and also to long-term memory access (Geary, 1993; Geary et al., 1999; McLean & Hitch, 1999; Swanson & Siegel, in press).

In studies of older children there is some support for the hypothesis that co-occurring RD and MD may share a common underlying deficit in retrieval from long-term or semantic memory: Children who are slow readers make more errors in retrieving math facts from memory than children who are neither RD nor MD and more than children who have MD, but not RD (Rasanen & Ahonen, 1995). Younger children with both RD and MD difficulties have poor counting knowledge. They treat counting as a rote activity, rather than having a solid understanding of the principles of counting, and these children also have difficulty on working memory tasks and on tests tapping retrieval of verbal semantic information from long-term memory (Geary et al., 1999). Other studies suggest that children who are impaired in both reading and math computations typically show more severe and pervasive disturbances of oral language than children who are only impaired in word recognition. Their difficulties reflect problems learning, retaining, and retrieving math facts, which are essential to precise calculation; these problems lead to pervasive difficulties with math. Thus, Jordan and Hanich (2000) found that children with both reading and mathematics difficulties showed problems in multiple domains of mathematical thinking. Working memory is generally more severely impaired in children with both RD and MD than in either alone (Swanson & Siegel, in press).

Specific math disability

The error-prone use of developmentally immature procedures and strategies in simple arithmetic, and perhaps in written arithmetic, may underlie the form of MD that is not related to RD (Geary, 1993). Data on the validity of this source of MD is less strong at the present time than data on memory-based deficits in math fact retrieval. At a younger age, these children may also have problems in counting and often make errors in the application of algorithms. Written problems involving carrying and borrowing are difficult, as is the learning of algorithms necessary to complete complex multiplication and division. There are no studies of which we are aware that link developmentally-immature and error-prone counting strategies in solving simple arithmetic problems in early childhood to later use of developmentally-immature procedures in solving more complex multidigit written arithmetic problems.

To return to our example of children with spina bifida and hydrocephalus who have profound difficulties with math, Barnes et al. (in press) showed that good readers with this form of brain injury made more procedural errors than age-matched controls, but similar numbers of math fact retrieval and visual-spatial errors. Furthermore, their procedural errors were similar to those of younger children who were matched in math ability with these older brain-injured children. In other words, the good readers with hydrocephalus made errors in written computation that were developmentally immature for their age, but not different in kind from younger children with no MD. These data are consistent with the hypothesis that children who are good readers but who are poor at math can have a procedural deficit that involves the application of developmentally immature algorithms for solving written computations.

Figure 7.Profiles of cognitive performance by children with only reading disability (RD), only math disability (MD), both RD and MD (RD-MD), attention-deficit hyperactivity disorder (ADHD) with no RD or MD, and no LD (NL). The profiles show differences in shape and level of performance suggesting that the groups are different with distinct patterns associated with RD, MD, and ADHD as well as areas of overlap suggesting comorbidity of RD, MD, and ADHD. The NL group is clearly different from all three groups with LD in shape and elevation.

Comparison of specific MD, RD, and RD-MD

Children with word-level RD, computational MD and no RD, and both RD and MD can be differentiated. Figure 7 compares these groups with a contrast group of children who only have problems with behavior, meeting diagnostic criteria for ADHD but with no evidence of LD. The contrast with ADHD is important as some of the hypothesized cognitive correlates of specific MD are also apparent in studies of children with ADHD (Barkley, 1997), though such studies rarely address the issue of comorbidity. A group with no LD (NL) is also included.

These groups are compared on variables that have been related to ADHD (attention, paired associate learning), both ADHD and MD (problem solving, concept formation), and RD (phonological awareness, rapid naming, and vocabulary). Figure 7 demonstrates the pervasive problems experienced by children who have specific or comorbid RD on measures of phonological awareness. In addition, whereas children with

MD had the most significant difficulties with concept formation, children with RD and MD, ADHD without LD, and only MD share in common difficulties on a problem solving measure. Finally, it is clear that children with both RD and MD are the most pervasively impaired, whereas children with no LD and only ADHD show much stronger performance in areas that involve language, working memory, and visual-motor integration. The group with only ADHD has problems with sustained attention (continuous performance test) and procedural learning, the latter representing strength in the group with only RD.

If Table 7 broke the three LD groups out by IQ-discrepancy/low achieving or by presence/absence of ADHD, the effect would be only on the level of the profiles, not the patterns. As the RD-MD example demonstrates, whenever a child has a disability in more than one area, their overall performance is lower. While IQ may also be lower, it is knowing that the child is disabled in more than one domain (including ADHD) that is critical. Thus, assessment of the academic (and behavioral) domains, and cognitive correlates, are the keys to understanding the disability—not level of IQ.

The differences in results across studies of LD and ADHD may well reflect variations in whether these domains are assessed. For example, children with only ADHD have small cognitive impairments relative to any child with LD. A study that combined children with only ADHD and both ADHD and RD-MD would show more significant cognitive impairments that may be attributed to ADHD if the comorbidity is not addressed. It is not surprising that studies of children defined as LD without specification of the academic domains that are impaired have not contributed much to research or practice.

What is particularly intriguing about the differences in the three LD subgroups (RD, MD, and RD-MD) comes from the possibility that the neural correlates are different. Functional imaging studies of children with RD and RD-MD reliably demonstrate aberrant activations involving the left temporoparietal areas (see Constitutional Factors below). Although there are presently no functional neuroimaging studies of children with math difficulties, studies of how math is represented in normal adults show that there are different neural correlates of precise calculation versus estimation (Dehaene, Spelke, Pinel, Stanescu, & Tsiukin, 1999). Precise estimation involves the inferior prefrontal cortex in the left hemisphere, as well as the left angular gyrus. These areas overlap substantially with those that mediate language functions. In contrast, estimation tasks showed bilateral activation in the inferior parietal lobes, which represent areas that overlap with spatial cognition. As many children with specific MD have been found to also have spatial cognition difficulties, this overlap in neural representation of estimation and spatial cognition may help explain why the spatial processing difficulties do not seem to bear a strong relationship to math abilities in these children, but are often as profound as the math difficulties themselves. Any cognitive task sensitive to how these areas of the brain function will be deficient in children with specific RD, but this does not mean that the cognitive deficits themselves are tightly linked.

Altogether, there is burgeoning evidence for the existence of a group defined by difficulty in learning and retrieving math facts from memory (RD-MD), and some evidence for the existence of a group that has difficulty learning math calculations because of procedural difficulties (MD only). There is little evidence for a separate subgroup with impairment in math concepts, but this possibility has not really been studied. In a sense, all children with disabilities in math probably have difficulty at some level with math concepts broadly conceived. The meaningfulness of this putative category of LD—math concepts disability— is not clear.

Written Expression

There is also research on disorders that involve written expression (Berninger & Graham, 1998; Graham & Harris, 2000). This is clearly an area of difficulty for many children. In some students, writing difficulties reflect an inability to spell, most closely associated with difficulties in word recognition skills (Rourke, 1993). Even some children with specific MD can have difficulty with handwriting, often because they commonly have impairments in their motor development. Their spelling errors, interestingly, are typically phonetically constrained, in contrast to children who have word recognition difficulties (Rourke, 1993). Once these two difficulties (spelling and motor skills) are taken into account, is there a subgroup of children whose difficulties are restricted to written expression? Here the classification research that is necessary to evaluate this hypothesis has not been completed, but there is some evidence for this possibility. In particular, some children have specific problems with handwriting and respond to prevention interventions (Graham, Harris, & Fink, 2000). Future research should target this possible subgroup in an effort to flush out the heterogeneity hypothesis.

Conclusions: Heterogeneity

There is support for the heterogeneity classification hypothesis. It is clear that there are at least two types of RD (word recognition, comprehension) and probably a third (reading fluency). In addition, there is evidence for a form of specific MD involving calculations. Children with both RD and MD have problems associated with either domain alone, reflecting more pervasive disruptions of language and working memory (Rourke, 1993; Swanson & Siegel, in press). This type of LD should be differentiated from specific RD and specific MD. Research is weakest for disorders of written expression.

Children with disorders in listening and speaking can be differentiated from children who have problems with reading and math. Although there is overlap, only about 50% of those who develop specific language disorders also develop reading and math disorders (Tomblin & Zhang, 1999). When a child with a speech and language disorder develops a reading or math problem, it is for the same reasons that a child with a reading and math problem and no disorder of oral language develops these difficulties. To illustrate in the area of word recognition skills, it is because the child does not develop adequate phonological awareness skills, has problems with rapid naming, and has deficient vocabulary and oral language comprehension skills. There is little evidence for meaningful dissociations of listening and reading comprehension when word recognition skills are adequate. Research in reading comprehension disabilities has largely proceeded from the assumption that the comprehension disability occurs in both reading and listening (with many researchers using oral language tasks to measure components of comprehension), and the data have amply demonstrated that when reading decoding and reading fluency are intact, the comprehension deficit is similar in written and oral language (e.g., Stothard & Hulme, 1992). Thus, many children who have specific problems with reading comprehension have parallel difficulties in listening, i.e., understanding oral language.

Why disorders of listening and speaking are included in the LD category is unclear, as there is a separate category for speech and language disorders and disorders of listening and speaking are not specific academic skills disorders. Dropping them from the LD category would increase the conceptual clarity of the LD category.

Table 2 provides a hypothetical classification of types of LD. Although the evidence for each of these types varies and there is undoubtedly overlap, there is support for each of these six types. Each of the academic domains representing the disorders can be measured and something is known about the cognitive correlates of each domain, except possibly for written expression. As such, a single definition for all these forms of LD seems less useful than a set of inclusionary definitions that specific the domain and its associated cognitive correlates. Such an approach would be more directly related to intervention and would facilitate communication.

Table 2. Comprehension in poor comprehenders with and without known brain pathology (Barnes & Dennis, 1996)

Exclusion Hypothesis

Most definitions of LD include an exclusion clause, which simply states that LD is not primarily the result of other conditions that can impede learning. These other conditions include mental deficiency; sensory disorders; emotional disturbance; cultural, social, and economic conditions; and inadequate instructional opportunities. Given the role of the exclusion element within definitions of LD, children identified as LD are often identified on the basis of what they are not, rather than what they are. This is unfortunate for three major reasons. First, by placing an emphasis on exclusion, the development of inclusionary characteristics that are linked to assessment and intervention is difficult. To illustrate, the 1977 operationalization of the federal definition suggests that RD be assessed with IQ and achievement tests; it does not specify the domains of reading and ignores the variation in components of reading that may represent differential treatment emphases. Second, an exclusionary definition is a negative definition that adds little to conceptual clarity and clearly constrains understanding the disorder to its fullest extent. Think of the difference in clarity when we identify a child as LD versus identify the child with a reading comprehension disability. Third, many of the conditions that are excluded as potential influences in LD are themselves possible factors that interfere with the development of those cognitive and linguistic skills that lead to the academic deficits that form the basis for LD (Lyon et al., 2001). Parents with reading problems, for example, may find it difficult to establish adequate home literacy practices because of the cumulative effects of their reading difficulties (Wadsworth et al., 2000).

It is reasonable to stipulate that children with mental deficiency and sensory disorders are excluded from classifications of LD. Separate categories exist and their treatment needs are different. This stipulation begs the question of how to differentiate mild mental deficiency and LD (Gresham, MacMillan, & Bocian, 1996), given the weakness of psychometric definitions. Other exclusions are even more difficult to justify. For example, where is the evidence suggesting that RD and MD are different in children who are anxious, depressed, or even psychotic? Recent longitudinal studies suggest that early achievement is causally related to and often precedes the development of behavioral problems, and the interventions that enhance academic achievement prevent behavioral difficulties (Kellam, Rebok, Mayer, Ialongo, & Kalodner, 1994; Onatsu-Arvilommi & Nurmi, 2000). Thornier are exclusions based on social, economic, and cultural disadvantage, and inadequate instruction. This differentiation is based on the presumption that constitutional factors are more relevant for children with LD than environmental factors. In addition to reviewing evidence for exclusion according to environmental factors, we will also review additional evidence on the role of constitutional factors in LD, raising again the question of how well either environmental or constitutional factors distinguish children with LD from low achievement commonly ascribed to environmental factors.

Social, Economic, and Cultural Disadvantage

A variety of factors related to the literacy environment in which a child develops are clearly related to the acquisition of academic skills. When optimal social and economic conditions are not present, the child is at a much higher risk for the development of an academic problem. In reading, a variety of factors have been studied, including print exposure, parental literacy levels, and reading to the child. All these factors are related to the development of reading skills (Adams, 1990) and probably to other academic skills as well. Recent qualitative studies (Hart & Risley, 1999) have provided graphic documentations of the differences in the language environment experienced by advantaged and disadvantaged children. For example, by the age of 5 years, economically advantaged children have vocabularies of approximately 500,000 words, while economically disadvantaged children have vocabularies of approximately 250,000 words (Hart & Risley, 1999). It is widely believed that these types of differences in language development have some (unspecified) effect on brain development, and they are certainly related to the development of proficiency in academic skills.

These types of factors impede oral language development. When oral language development is affected, a variety of language skills are at risk, including those related to the development of word recognition and reading comprehension skills. In a series of longitudinal studies, Whitehurst and Lonigan (1998) provided excellent documentation of the relationship of different oral language skills and the acquisition of reading ability. In evaluating children entering Head Start programs at age 3, Whitehurst and Lonigan (1998) found that skills related to knowledge of the alphabet and word structures were closely tied to reading success in kindergarten, Grade 1, and Grade 2. More general oral language skills that involved vocabulary, language comprehension, and exposure to language through literature and oral reading were related to the development of reading comprehension skills, particularly in Grades 2 and 3.

These results are particularly striking when national evaluations of Head Start programs are examined

(Whitehurst & Massetti, in press). These studies have shown that, on average, children who graduate from Head Start programs enter kindergarten knowing one letter of the alphabet. Interviews with teachers showed that they were often discouraged from engaging in activities that promoted understanding of the alphabetic structure of the language because such activities were not viewed as "developmentally appropriate." Nonetheless, when children in Head Start programs were provided with these sorts of activities, higher literacy levels were apparent (Whitehurst & Lonigan, 1998).

What is important in these examples is the illustration that environmental factors influence the development of oral language skills that are known to affect beginning (and later) reading skills. Interventions that address the early development of these skills seem to promote success in reading. Such findings are also apparent in evaluative studies of Title I Programs as well as intervention studies in which alphabetic forms of instruction have been shown to be advantageous for economically disadvantaged children (Foorman et al., 1998; National Reading Panel, 2000). Thus, the mechanisms and practices that promote reading success in advantaged populations appear to be similar to those that promote reading success and failure in disadvantaged populations. There is little evidence that the phenotypic representation of RD varies according to socioeconomic status. Children at all levels of socioeconomic status appear to have reading problems predominantly (but not exclusively) because of word-level difficulties apparent in the beginning stages of reading development (Foorman et al., 1998; Wood, Flowers, Buchsbaum, & Tallal, 1991).

As Kavale (1988) and Lyon (1996) pointed out, the basis for excluding disadvantaged children from the LD category has more to do with how children are served than with empirical evidence demonstrating that characteristics of reading failure are different in LD in economically disadvantaged groups. Indeed, Kavale (1988) suggested that arguments usually point to the fact that "the culturally disadvantaged child is well served by various federally funded title programs, but these are usually mandated under guidelines and revisions different from special education. Specifically, the emphasis is on compensatory education while special education programs function as remedial programs" (p. 195). This has the effect of eliminating economically disadvantaged children from special education services, with the exception of categories related to mental deficiency and emotional disturbance; economically disadvantaged children are disproportionably represented in these special education categories. As Kavale stated,

> since culturally disadvantaged children have been shown to exhibit the behavioral characteristics included as primary traits in definitions of LD, it is difficult to determine why the culturally disadvantaged group is categorically excluded from the LD classification. Yet, children from lower SES levels with LD-type behaviors have little chance for receiving LD diagnoses and treatment with an increased likelihood of being labeled retarded in spite of the fact that LD and ED groups are not clearly identifiable as separate entities. (p. 205)

There is little empirical evidence supporting the exclusion of economically disadvantaged children from special education services as a valid classification practice. The exclusion is a policy decision that represents a desire to clearly separate funds dedicated to special education and compensatory education. We have, essentially, a two-tier service delivery system for children with academic difficulties, where advantaged children are designated as LD and served through remedial classes that are questionably effective (Lyon et al., 2001). In contrast, children who qualify for free lunches served under Title I often receive compensatory education programs, which in some studies appear to be effective (Slavin, Karweit, & Madden, 1989). Our concept of LD, however, must not hinge on policy issues. Here there is little basis for distinguishing types of poor achievement according to putative causes, since the phenotypic manifestations seem to be similar across levels of socioeconomic status.

Instruction

Virtually any definition of LD excludes children from consideration if their learning problems are primarily a product of inadequate instruction. Of all the different assumptions in the concept of LD, this assumption is the least frequently examined and perhaps the most important. Some would interpret the exclusion to indicate that children who profit from instruction do not have a biologically based disorder. The functional imaging studies reviewed below suggest that this is hardly the case and that instruction is necessary to establish the neural networks that support reading. Keep in mind that no child is born as a reader; all children are taught to read. Written language is scaffolded upon our natural capacities for developing oral language (Lukatela & Turvey, 1998). It may be that there are differences in brain function that make some children more refractory to intervention than others, but we do not presently have data that would indicate that this is the case.

Another problem with the inadequate instruction exclusion is that it presumes that the field has a good understanding of what constitutes adequate instruction. At the time the federal definition was adopted, this was not case. Recent consensus reports (National Reading Panel, 2000; Snow et al., 1998) make it clear that we do know a lot about teaching children to read. Given what we know, consideration of the students' response to well designed and well implemented early intervention as well as remediation programs may need to become part of the definition of LD. Why should the complex identification criteria and expensive due process procedures of special education be used before an attempt is made to provide a powerful intervention early in the child's development? A child's failure to respond to intervention may be the best way to operationalize the notion of adequate instruction. While a child's failure to respond to appropriate instruction is a very strong indication of a disability, the cognitive problems associated with their LD parallel those exhibited by children who do not respond to inadequate instruction. The two types of children are equally disabled and there is no evidence that there are differences at a neurological level, prior to intervention or in terms of their intervention needs, that would make them different. For children with mental deficiency, sensory disorders, and emotional disturbance, there are other classifications in IDEA that can lead to services. For the child who is deemed culturally, economically, or socially disadvantaged, compensation education programs are available. What is there for the child who develops academic difficulties because of poor instruction? Excluding children on the basis of inadequate instruction does not seem a reasonable practice.

Constitutional Factors

Approaching the exclusion hypothesis from the perspective of classification research shows little evidence supporting exclusions based on emotional disturbance; social, cultural, and economic disadvantage; or inadequate instruction. This reflects the difficulties of differentiating forms of low achievement that are presumably specific or unexpected from those than can be attributed to other causes, where low achievement is expected. Related to this hypothesis is another source of data that is frequently invoked in explaining unexpected low achievement. That is the notion that unexpected LD is due to constitutional factors that are intrinsic to the child. In the current federal definition of LD, the intrinsic/neurological component is implicit in the use of terms like "basic psychological processes." In other contemporary non-federal definitions the concept is explicitly stated. For example, the definition of LD proposed by the National Joint Committee on Learning Disabilities (1988) states: "these disorders [LD] are intrinsic to the individual, presumed to be due to central nervous system damage, and may occur across the life span" (p.1).

Neurobiological factors do not represent formal classification hypotheses in the sense that they are used to identify students with LD. They do represent components that can be tested for validity purposes. If children with unexpected low achievement differ from children in whom achievement is expected on constitutional factors, then this might support the hypothesis that expected and unexpected low achievement should be differentiated.

It has long been assumed that neurobiological factors were the basis of LD, reflecting its conceptual origins in the notion of organically based behavior disorders (Doris, 1993; Rutter, 1982; Satz & Fletcher, 1980). Neurobiological (constitutional) dysfunction was inferred from what was then known about the linguistic, cognitive, academic, and behavioral characteristics of adults with documented brain injury or lesion. As the field progressed, definitions of LD continued to attribute the disabilities in learning to intrinsic (neurological) rather than extrinsic (e.g., environmental, instructional) causes, even though there was no objective way to adequately assess the presence of putative brain damage or dysfunction. These assumptions of constitutional etiology were buttressed by associations of a variety of indirect indices of neurological dysfunction and LD. These indirect indices included observations of perceptual-motor problems (i.e., difficulty copying geometric figures), paraclassical or "soft" neurological signs (e.g., gross motor clumsiness, fine motor incoordination), and anomalies on electrophysiological measures, such as an electroencephalogram (Taylor & Fletcher, 1983). Even at the time, the lack of specificity of these observations with either LD or neurological integrity was widely acknowledged (Satz & Fletcher, 1980). Nevertheless, the neurobiological deficits were presumed to be selective rather than diffuse, resulting in specific difficulties processing linguistic, visual, and motor information critical to academic learning without concomitant loss of general intellectual functions.

Over the past two decades, some evidence, varying widely in methodological quality, has been obtained from investigations designed to identify, more directly, the neurological basis for LD and particularly RD. For example, data derived from postmortem studies performed on dyslexic adults and structural neuroimaging studies with children and adults have indicated that some individuals with RD are characterized by differences in the size of specific brain structures (e.g., planum temporale) and in the presence of specific neuroanatomical anomalies (e.g., ectopias) (Filipek, 1996; Galaburda, 1993). Structural imaging studies reliably show that people with RD have a smaller left hemisphere, or less asymmetric hemispheres. Both the autopsy and structural imaging studies have been confounded by subject selection problems, failure to account for comorbid neurological disease (e.g., seizures) and other variables (e.g., handedness). Interpretation of the structural imaging studies has been impeded by the use of different neuroimaging methods and data analytic techniques, as well as difficulties replicating the findings of these studies (Filipek, 1996; Shaywitz et al., 2000).

More recently, research using different types of functional neuroimaging methods to measure brain activation in response to visual, linguistic, and reading tasks among skilled and unskilled readers indicates systematic and selective brain activity in several left hemisphere neural systems subserved by the basal surface of the temporal lobe, the middle temporal gyrus, the temporoparietal region, and the inferior frontal region. Converging evidence from a range of functional imaging methods used in studies with both good and poor readers indicate that a network of brain areas is involved in the ability to recognize words accurately, and that adults and children with RD manifest different patterns of activation in these areas when compared with skilled readers (Shaywitz et al., 2000).

A critical question that has been raised by the functional neuroimaging studies of those with LD in reading is whether the patterns seen in these individuals with RD are compensatory in nature ("compensatory hypothesis") or reflect the failure of the environment and/or instruction to impact the brain in a manner necessary to form the neural networks that support word recognition. Thus, the pattern in RD children may be similar to that seen in a young child who has not learned to read and may change by virtue of development, instruction, or even intervention ("normalization hypothesis"). Given this possibility, functional neuroimaging studies may provide an example of how brain and environment interact in forming neural networks for complex behaviors. Such studies are feasible and investigations that combine neuroimaging and reading intervention studies are currently being completed.

Figure 8. Individual activation maps from a 10-year-old child who was experiencing serious difficulties in learning to read, before and after an intense phonologically-based intervention. Activation maps were obtained using a pseudoword rhyme-matching task. The child showed dramatic improvement in phonological decoding skills after 8 weeks (80 hours) of enrollment in the program, analogous to that reported by Torgesen et al. (2001). Note the dramatic increase in the activation of the left temporoparietal regions.

Figure 8 provides an example from a pilot study in which functional neuroimaging studies were performed using magnetic source imaging while a child read words. The imaging studies occurred before and after approximately 60 hours of intense intervention (over 8 weeks) in which the child, who was 10 years old with severe RD, showed significant improvement in word reading ability into the average range. The top part of the figure shows the standard brain activation pattern characterized by activity predominantly in the temporoparietal regions of the right hemisphere. After intervention, the pattern shifts to predominant activation involving the homologous areas of the left hemisphere, an activation pattern typical of non-disabled readers. Thus, these results are more consistent with the normalization hypothesis than the compensatory hypothesis.

The preliminary data from these types of studies suggest different conceptualizations of the role of constitutional factors in LD. The view that is emerging suggests that neural systems develop and are deployed for specific behaviors through the interaction of brain and environment (including instruction) as opposed to representing fixed properties of the nervous system that inherently limit learning potential. As such, the concept of LD retains the optimism that was intended with its inception.

This interaction perspective is also supported by genetic studies of individuals with RD. It has long been known that reading problems reoccur across family generations, with a risk in the offspring of a parent with RD 8 times higher than in the general population. Multiple genes are most likely involved, with similar modes of transmission in dyslexic and non-dyslexic families. Linkage studies implicate markers on chromosomes 1, 2, 6, and 15. However, genetic factors account for only about half of the variability in reading skills, which means that the environment has a significant influence on reading outcomes. This also suggests that what is inherited is a susceptibility for RD that may manifest itself given specific interactions, or lack thereof, with the environment. For example, parents who read poorly may be less likely to read to their children. As such, the quality of reading instruction provided in the school may be more critical for children when there is a family history of poor reading giving rise to limited environmental-instructional interactions in the home (Olson, Forsberg, Gayan, & DeFries, 1999; Pennington, 1999; Wadsworth et al., 2000).

Conclusions: Exclusionary Criteria

There is little evidence that children excluded from LD classifications due to emotional disturbance; social, economic, and cultural disadvantage; or instructional history are meaningfully different from those included as LD. In particular, none of these criteria provide robust differentiations of expected and unexpected low achievement. The notion that expected and unexpected low achievement reflects variation in cognitive and behavioral correlates, prognosis, response to instruction, or even a broad range of neurobiological factors, does not have strong validity. This does not mean that the concept of LD is not valid or that the exclusions should not be used, particularly since many children can be served under other categories in IDEA or other approaches to providing services (e.g., compensatory education). There may well be needs outside the academic area that are better addressed through identification for other categories or programs. Exclusions due to inadequate instruction are not justifiable as lack of instruction can essentially cause LD. The exclusions must be seen as policy-based determinations to facilitate service delivery and avoid commingling of facts, not as classification factors that have strong validity.

Future Directions for Classifications of LD

In this paper, we have reviewed federal and non-federal definitions of LD, pointing out that these definitions embed hypothetical classifications at three levels: IQ discrepancy, heterogeneity, and exclusion. We also evaluated the evidence for the hypothesis that LD can be related to constitutional factors, showing that environmental factors must be accounted for in explaining not only why a child develops LD, but also the role of instruction. Throughout the paper, we highlighted some alternative approaches to classification, reflecting different ways of thinking about LD. We focused specifically on the value of inclusionary definitions that identify specific forms of LD, leading to specific (and less time-consuming) identification practices that we believe are directly linked to intervention. We suggested a hypothetical reorganization of the types of LD identified in the 1977 operationalization of the federal definition of LD. We recognized that exclusions, with the exception of inadequate instruction, largely reflect that there are other ways of serving children including different categories in IDEA and other services, such as compensatory education. There is little evidence that children meeting these exclusionary criteria have different instructional needs or respond differently to intervention.

Kavale and Forness (2000) presented an approach to the classification, definition, and identification of LD that in many respects is the antithesis of what would be recommended based on the research reviewed in this report. It begins with the acceptance of unexpected low achievement at the first level and the notion that discrepancy sets apart a specific form of LD as a necessary but not sufficient criterion at the first level of identification. The approach recognizes the heterogeneity of LD at the second level, tying LD to achievement deficiencies in language, reading, writing, and math. These deficiencies presume the presence of IQ discrepancy. At levels III and IV, issues related to learning processes are added. At level V, children are excluded because of sensory impairment, mental deficiency, emotional disturbance, social and cultural disadvantage, and inadequate instruction.

This approach to classification hinges on the validity of IQ discrepancy as demarcating a specific form of LD that is differentiated from low achievement. The evidence reviewed in this paper shows that IQ-discrepant and low achieving groups overlap substantially in cognitive characteristics and show little difference in response to intervention and long-term outcome. Similar problems affect the use of exclusionary criteria. Consider children who have (a) an IQ discrepancy, (b) problems in reading, (c) processing difficulties, and (d) who do not meet any of the exclusionary criteria. How are they meaningfully different or have different instructional needs from children with (a) through (c), but who meet exclusionary criteria and are therefore not defined as LD? There is little evidence that this would be the case, even if the hypothesis was tested only within IQ-discrepant children. There is really no hypothesis to test, as there is no basis for imagining how such subgroups could differ if the sorting was based solely on the exclusionary criteria. Even if one argued that exclusion would be infrequent because of all the prior levels of identification, the evidence in this paper does not support the hypothesis that children excluded as non-LD are meaningfully different from those who make the cut.

We do not mean to indicate that federal and non-federal definitions that have been used to the present have not had utility. On the positive side, the evolution of the current federal definition in IDEA has successfully served as a rallying point for special interest groups and for increased funding for special education programs. The current omnibus federal definition has served well as a galvanizing force for advocacy groups in their quest to obtain funding and secured educational services support for children with LD. Current (and historical) classifications and the resultant definitions of LD should be conceptualized as hypotheses that require rigorous, ongoing evaluation. The review of evidence in this paper shows that the classifications have become obsolete and should be revised, especially if the goal is to guide and reform instruction.

A major problem is the notion that low achievement in LD is unexpected, leading to a focus on exclusion. The accumulation of research over the past 30 years shows that low achievement is expected and suggests a focus on identifying the factors responsible for poor achievement in every child. Such a shift would suggest the need to develop inclusionary definitions that build upon the cumulative research base on LD. The move from exclusionary to inclusionary definitions is the first of many steps.

Psychometric Approaches Are Limited

In addition to the need for inclusionary definitions, we must recognize that an approach to identification based solely on test scores is not likely to be reliable and begs the question of where to put the cut-score. Achievement test scores are continuous and largely normally distributed. The tests used to measure these domains have measurement error. Any attempt to set a cut-point will lead to instability around the cut-point as scores fluctuate around the point with repeat testing, even for a decision as straightforward as demarcating low achievement. This fluctuation is not a problem of repeat testing, nor is it a matter of selecting the ideal cut-score. The problem stems from the fact that no single score can perfectly capture a student's ability in a single domain. There is always measurement error. Fluctuation will also vary across tests, depending in part on the cut-score, as tests vary in their precision at various ranges of the ability scale. This problem is more significant as the cut-point moves from the center of the distribution.

A second problem with the typical use of cut-scores concerns their arbitrary nature. A cut-point on a norm-referenced test is an arbitrary, relative standard of performance. The arbitrariness of the standard does not mean that a cut-point does not indicate a problem. Rather, arbitrariness reflects the meaninglessness of distinctions between, for example, the 15th and 20th percentile (or the 20th and 21st percentile). The problem with arbitrariness is not so much with the use of norm-referenced tests for establishing cut-points, but reflects difficulties inherent in any approach that would make critical decisions based on a single indicator. A single assessment at a single point in time is not psychometrically adequate for deciding placement. The flexibility in IDEA that allows interdisciplinary teams to go beyond test scores and encourages clinical judgment is necessary because of these issues. But the basis for clinical judgment should include performance on psychometric tests that involve achievement and cognitive performance. Inclusionary definitions based on patterns on these types of tests may be especially useful.

IQ Tests Are Not Needed

Such an approach would dramatically reduce the reliance on IQ tests for the identification of LD. Although there may be a role for IQ tests in determining mental deficiency, even here the more important concept is adaptive behavior, and there are difficulties establishing the upper range that distinguishes mental deficiency from LD (MacMillan, Siperstein, & Gresham, 1996).

The problems that we observed above in setting cut-points also apply to IQ distributions. There is no natural subdivision that demarcates mental deficiency from LD. Even with the stipulation of mental deficiency, there is no need to give every child referred for special education an IQ test. For LD, the information has limited relevance, particularly for intervention. The concept of IQ as it is applied to LD is outmoded and reflects an obsolete practice. The use of IQ tests reflects a focus on compliance as opposed to results that must shift if placement in special education as LD is to benefit the person so designated. IQ tests do not measure aptitude for learning or provide an index of response to intervention. The processes that contribute to performance on an IQ test may well be an outcome of the same processes that led to the LD. Dropping IQ from the LD definition would shift the focus to achievement/cognitive processes and also result in more efficient, less expensive evaluations.

"Slow Learner" Is Not a Useful Concept

Related to the issue of the obsolete role of IQ for LD is the notion of the *slow learner*, or *garden-variety* poor learner. These terms are also used to refer to children with low achievement at levels consistent with their IQ. There are clearly children who have impairments in multiple cognitive and academic domains who obtain lower scores on IQ tests. Many of these children represent what we described earlier as the comorbid RD-MD group.

Although it is commonly assumed that IQ is an indicator of the slow learner, this does not appear to be the case. It is difficult to identify an IQ cut-point, even in the non-mentally deficient range, that would differentiate specific LD and garden-variety LD. IQ scores do not reliably differentiate children with different types of LD. To illustrate, McFadden (1990) completed a cluster analysis to determine whether level of IQ was associated with different types of LD. McFadden found that (1) children with IQs between 70 and 80 were generally represented in all clusters of children with learning disabilities; (2) many children with low IQs exhibited similar patterns of cognitive difficulties relative to children defined as having learning disabilities by discrepancy criteria; (3) although a WISC full scale IQ cut-off of 80 reduced the number of children with low IQs in learning disabilities clusters, several subtypes still contained children with approximately 20 percent lower IQs; and (4) children with low IQs were apparent in clusters of children with learning disabilities and, within such clusters, differences occurred in level but not shape. These results question the validity of differentiating learning disabilities according to IQ cutoffs of 80 and above, but do not identify appropriate cutoffs (if any).

In another cluster analytic study, Morris et al. (1998) were able to distinguish children with specific RD who had cognitive problems relatively restricted to the phonological domain from those who had more generalized difficulties in multiple cognitive domains (e.g., vocabulary, speech production, attention). On average, children with non-specific RD had lower scores on IQ tests than children with specific RD, but IQ ranged considerably within each subtype. The differences between specific and non-specific subtypes were most reliably indexed by the child's vocabulary development and could be understood as the consequences of the child's poor language development, which in turn produced lowered IQ scores.

The notion that low achievement is expected in garden-variety RD and unexpected in specific RD is also specious. The basis for reading difficulties was associated with phonological processing in all subtypes with word reading problems. Some specific subtypes read as poorly as the non-specific subtypes, but the groups did not differ qualitatively in language characteristics related to reading (i.e., in the phonological domain). The garden-variety group may well have a poorer prognosis and need different types of instruction. The garden-variety group may even show different neurological characteristics. But would we really want to restrict our concept of LD or eligibility for services to children with specific types? Schools are interested in serving the lowest achievers as these are most difficult to teach (MacMillan et al., 1996).

Research on children with LD has not progressed to the point where we can say definitively that children with specific and garden-variety subtypes need different interventions, have different prognoses, or respond differently to treatment. This reflects in part the preoccupation with concepts of LD based on unexpectedness, IQ-discrepancy, and anxiety over the role of underachievement in LD. The consolidating issue is that the concept of underachievement and the linking of LD to an academic deficiency (e.g., reading, math) are *necessary* to the concept of LD. They are not *sufficient* and it is essential to include the concept of process (e.g., language, perceptual skills) as necessary to the concept of LD (Kavale & Forness, 2000). It is also essential to drop notions of "potential," "ability," and their operationalization in measures of IQ and to move towards attributes or components that are measurable and linked to intervention. Thus, we would move the concept of LD from a disorder that is unexpected because of discrepancies between ability potential and achievement to one in which underachievement is expected because of impairment of key cognitive processes. These processes are measurable and can be directly linked to intervention.

Response to Intervention Is Important

It is essential to introduce the student's response to well-designed instruction and remediation programs as a major component of the identification of LD. This introduction should be made in the context of early identification and prevention programs that are seen as fundamental to general and special education. Children who do not benefit from early and intensive interventions will require even more powerful remediation programs as well as educational accommodations as they proceed through their schooling. The information on how well the child responded operationalizes the "inadequate instruction" component and those who do not respond to increasingly intense interventions may indeed be disabled. In addition, continuous monitoring of progress will be helpful not only for instructional planning, but also for identifying those who do not respond to adequate instruction (Fuchs & Fuchs, 1998).

Consensus Process

To do justice to the need for a classification of LD that yields inclusionary definitions with the features we have identified as being desirable, we call for the development of a consensus process. As part of this process, the relevant federal agencies responsible for research and practice involving people with LD should work together to synthesize the available research. The principles and goals of a new overarching classification should be explicitly articulated, with specification of boundaries and overlaps with other classifications of childhood disorders (e.g., mental deficiency, emotional disorders). The possibility of comorbid associations should be incorporated. Working groups could be assembled to formulate definitions of different types of LD. Definitions should be formulated only for those types of LD where there is clear evidence of their nature and correlates. The classification and definitions should be treated as hypotheses. Research to evaluate the resultant classifications is desirable and should be supported. Plans to periodically update and revise the definitions should be made. Such a process would hopefully permit the development of specific procedures for identifying different types of LD that are efficient, that do not waste resources, and that lead to specific interventions.

Thus, we propose careful assessment of academic skills and their cognitive correlates as part of the implementation of inclusionary definitions. These assessments should be completed to address prevention/intervention needs with a goal of evaluating the instructional needs of the child. Adding continuous monitoring of progress and response to intervention as considerations in this process may go a long way towards the ultimate goal of helping as many children as possible master academic skills and return to regular education. A consensus process would help ensure that the last available evidence from research and the best available presentations were marked in a re-formulation of the federal classification of LD. Any changes must take into account the need for improved teacher preparation in general and special education, especially if the federal classification is changed.

Learning Disabilities Are Real Phenomena

Some researchers have confused the IQ-discrepancy hypothesis with the concept of LD, which is not appropriate. For example, Aaron (1997) stated that "when the discrepancy formula disappears from the educational science, so will the concept of LD" (p. 489). Similarly, Kavale and Forness (1994) stated that "… the notion of discrepancy … has led to a confounding … most clearly seen in the suggestion that there are more similarities than differences between LD and low achieving students. Such a suggestion calls into question the very notion of LD" (p. 43).

This conceptualization could be shown to be unreliable and invalid with no consequences for the validity of the concept of LD. This is clearly indicated in Figure 7, which shows that children with RD, MD, and RD-MD can be differentiated from those with no LD, even when ADHD is involved. Patterns of performance differentiate types of LD, while both level and patterns differentiate those with and without LD. Children identified with either an IQ discrepancy or LD definition *are* disabled, need to be identified, and respond similarly to appropriate educational interventions. What is being questioned is the validity of *classifications* of LD based on the presence or absence of IQ discrepancy and exclusion, not the reality of LD. Both definitions validly identify LD with or without the exclusions. As a classification, coverage, reliability, and validity are not adequate.

When the original federal definition of LD was proposed, there was little research that supported the discrepancy, heterogeneity, and exclusionary components of LD classifications. Since then research has accumulated suggesting that the discrepancy and exclusion components have (at best) weak validity and may be harmful and represent an obstacle to effective intervention. The goal should be to close the achievement gap for students identified as LD relative to their peers. Unfortunately, this happens all too infrequently by virtue of placing students with LD in special education (Lyon et al., 2001). One part of the solution is to revise the federal definition of LD and develop new classifications that are linked to research. New definition and identification practices will emerge, so that those who serve children with LD can focus on early identification, prevention, and effective remedial strategies. Eligibility and compliance presently consume excessive fiscal and emotional resources; this consumption should be redirected to intervention and special education should be re-oriented towards results, which means truly remediating children and returning them to the educational mainstream. These are the ultimate purposes of classifying a student as LD and the reasons that such classifications were developed. Such purposes must guide the reworking of the federal classification of LD essential to ensuring that all children can learn and reach their full potential in our society.

References

Aaron, P. G. (1997). The impending demise of the discrepancy formula. *Review of Educational Research*, *67*, 461–502.

Ackerman, P. T. & Dykman, R. A. (1995). Reading-disabled students with and without comorbid arithmetic disability. *Developmental Neuropsychology, 11*, 351–371.

Adams, M. J. (1990). *Beginning to read*. Cambridge, MA: MIT Press.

Algozzine, B., Ysseldyke, J. E., & McGue, M. (1995). Differentiating low-achieving students: Thoughts on setting the record straight. *Learning Disabilities Research*, *10*, 140–144.

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, *44*, 75–106.

Badian, N. A., (1999). Reading disability defined as a discrepancy between listening and reading comprehension: A longitudinal study of stability, gender differences, and prevalence. *Journal of Learning Disabilities*, *32*, 138–148.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*, 65.

Barnes, M. A., & Dennis, M. (1992). Reading in children and adolescents after early onset hydrocephalus and in normally developing age peers: Phonological analysis, word recognition, word comprehension, and passage comprehension skills. *Journal of Pediatric Psychology*, *17*, 445–465.

Barnes, M. A., & Dennis, M. (1996). Reading comprehension deficits arise from diverse sources: Evidence from readers with and without developmental brain pathology. In C. Cornoldi and J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 251–278). Mahwah, NJ: Lawrence Erlbaum Assoc.

Barnes, M. A., Dennis, M., & Wilkinson, M. (1999). Reading after closed head injury in childhood: Effects on accuracy, fluency, and comprehension. *Developmental Neuropsychology*, *15*, 1–24.

Barnes, M. A., Pengelly, S., Dennis, M., Wilkinson, M., Rogers, T., & Faulkner, H. (in press). Mathematics skills in good readers with hydrocephalus. *Journal of the International Neuropsychological Society*.

Berninger, V., & Graham, S. (1998). Language by hand: A synthesis of a decade of research in handwriting. *Handwriting Review*, *12*, 11–25.

Bisanz, J., Morrison, F. J., & Dunn, M. (1995). Effects of age and schooling on the acquisition of elementary quantitative skills. *Developmental Psychology, 31*, 221–236.

Cain, K., & Oakhill, J. V. (1999). Inference making and its relation to comprehension failure in young children. *Reading and Writing: An Interdisciplinary Journal, 11*, 489–503.

Cain, K., Oakhill, J. V., & Bryant, P. (2000). Phonological skills and comprehension failures: A test of the phonological processing deficits hypothesis. *Reading and Writing, 13*, 31–56.

Cataldo, M. G., & Cornoldi, C. (1998). Self-monitoring in poor and good reading comprehenders and their use of strategy. *British Journal of Developmental Psychology*, *16*, 155–165.

Clements, S. D. (1966). *Minimal brain dysfunction in children*. NINDB Monograph No. 3. Washington DC: U.S. Department of Health, Education, and Welfare.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, *7*, 249–253.

Cornoldi, C., DeBeni, R., & Pazzaglia, F. (1996). Profiles of reading comprehension difficulties: An analysis of single cases. In C. Cornoldi and J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 113–136). Mahwah, NJ: Lawrence Erlbaum Assoc.

Critchley, M. (1970). *The dyslexic child*. Springfield, IL: Charles C. Thomas.

Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. *American Educator, 4*, 8–15.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42.

Dehaene, S. (1999). Fitting two languages into one brain. *Brain, 122*, 2207–2208.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsiukin, S. (1999). Sources of mathematical thinking: Behavioral and brain-injury evidence. *Science*, *284*, 970–974.

Doris, J. (1993). Defining learning disabilities: A history of the search for consensus. In G. R. Lyon, D. B. Gray, J. F. Cavanaugh, & N. A. Krasnegor (Eds.), *Better understanding learning disabilities* (pp. 97–116). Baltimore: Paul H. Brookes.

Filipek, P. (1996). Structural variations in measures in the developmental disorders. In R. Thatcher, G. Lyon, J. Rumsey, N. Krasnegor (Eds.), *Developmental neuroimaging: Mapping the development of brain and behavior* (pp. 169–186). San Diego, CA: Academic Press.

Fletcher, J. M. (1985). Memor for verbal and nonverbal stimuli in learning disability subgroups:

Analyis by selective reminding. *Journal of Experimental Child Psychology, 40*, 244–259.

Fletcher, J. M., Foorman, B. R., Shaywitz, S. E., & Shaywitz, B. A. (1999). Conceptual and methodological issues in dyslexia research: A lesson for developmental disorders. In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders* (pp. 271–306). Cambridge, MA: MIT Press.

Fletcher, J. M., Francis, D. J., Rourke, B. P., Shaywitz, B. A., & Shaywitz, S. E. (1993). Classification of learning disabilities: Relationships with other childhood disorders. In G. R. Lyon, D. Gray, J. Kavanagh, & N. Krasnegor (Eds.), *Better understanding learning disabilities* (pp. 27–55). New York: Paul H. Brookes.

Fletcher, J. M., Francis, D. J., Shaywitz, S. E., Lyon, G. R., Foorman, B. R., Stuebing, K. K., et al. (1998). Intelligent testing and the discrepancy model for children with learning disabilities. *Learning Disabilities Research & Practice*, *13*, 186–203.

Fletcher, J. M., Francis, D. J., Stuebing, K. K., Shaywitz, B. A., Shaywitz, S. E., Shankweiler, D. P., et al. (1996). Conceptual and methodological issues in construct definition. In G. R. Lyon and N. A. Krasnegor (Eds.), *Attention, memory, and executive functions* (pp. 17–42). Baltimore: Paul H. Brookes.

Foorman, B. R., Francis, D. J., Beeler, T., Winikates, D., & Fletcher, J. M. (1997). Early interventions for children with reading problems: Study designs and preliminary findings. *Learning Disabilities*, *8*, 63–71.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, *90,* 37–55.

Francis, D. J., Fletcher, J. M., Shaywitz, B. A., Shaywitz, S. E., & Rourke, B. P. (1996). Defining learning and language disabilities: Conceptual and psychometric issues with the use of IQ tests. *Language, Speech, and Hearing Services in Schools*, *27*, 132–143.

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal individual growth curves analysis. *Journal of Educational Psychology*, *88*, 3–17.

Frankenberger, W., & Fronzaglio, K. (1991). A review of states' criteria for identifying children with learning disabilities. *Journal of Learning Disabilities*, *24*, 495–500.

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Lipsey, M. E. (2000a). Reading differences between low-achieving students with and without learning disabilities: A meta-analysis. In R. Gersten, E. P. Schiller, & S. Vaughn, *Contemporary special education research* (pp. 105–136). Mahwah, NJ: Lawrence Erlbaum Assoc.

Fuchs, D., Fuchs, L. S., Mathes, P. G., Lipsey, M. E., & Eaton, S. (2000b). A meta-analysis of reading differences between underachievers with and without the disability label: A brief report. *Learning Disabilities*, *10*, 1–4.

Galaburda, A. M. (1993). The planum temporale. *Archives of Neurology*, *50*, 457.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin*, *114*, 345–362.

Geary, D. C. (1994). *Children's mathematical development: Research and practical applications*. Washington, DC: American Psychological Association.

Geary, D. C., Bow-Thomas, C. C., & Yao, Y. (1992). Counting knowledge and skill in cognitive addition: A comparison of normal and mathematically disabled children. *Journal of Experimental Child Psychology*, *54*, 362–391.

Geary, D. C., Brown, S. C., & Samaranayake, V. A. (1991). Cognitive addition: A short longitudinal study of strategy choice and speed-of-processing differences in normal and mathematically disabled children. *Developmental Psychology*, *27*, 787–797.

Geary, D. C., Fan, L., & Bow-Thomas, C. C. (1992). Numerical cognition: Loci of ability differences comparing children from China and the United States. *Psychological Science, 3*, 180–185.

Geary, D. C., Hoard, M. K., & Hamson, C. O. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology*, *74,* 213–239.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Ginsburg, H. P., Klein, A., & Starkey, P. (1998). The development of children's mathematical thinking: Connecting research with practice. In W. Damon (Series Ed.), I. E. Siegel, & K. A. Renninger (Vol. Eds.), *Handbook of child psychology: Vol. 4. Child psychology in practice* (5th ed., pp. 401–476). New York: John Wiley & Sons.

Goldman, S. R., Pellegrino, J. W., & Mertz, D. L. (1988). Extended practice of basic addition facts: Strategy changes in learning-disabled students. *Cognition and Instruction, 5,* 223–265.

Graham, S., & Harris, K. R. (2000). Helping children who experience reading difficulties: Prevention and intervention. In L. Baker, J. Dreher, & J. Guthrie (Eds.), *Engaging young readers: Promoting achievement and motivation* (pp. 43–67). New York: Guilford Press.

Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write: Treatment of handwriting problems in beginning writers. *Journal of Educational Psychology*, *92*, 620–633.

Gresham, F. M., MacMillan, D. L., & Bocian, K. M. (1996). Learning disabilities, low achievement, and mild mental retardation: More alike than different? *Journal of Learning Disabilities*, *29*, 570–581.

Hart, B., & Risley, T. R. (1999). *Meaningful differences in the everyday experience of young American children*. Baltimore: Paul H. Brookes.

Hatcher, P., & Hulme, C. (1999). Phonemes, rhymes, and intelligence as predictors of children's responsiveness to remedial reading instruction. *Journal of Experimental Child Psychology*, *72,* 130–153.

Hoskyn, M., & Swanson, H. L (2000). Cognitive processing of low achievers and children with reading disabilities: A selective meta-analytic review of the published literature. *The School Psychology Review, 29*, 102–119.

Jordan, N. C., & Hanich, L. B. (2000). Mathematical thinking in second-grade children with different forms of LD. *Journal of Learning Disabilities*, *33*, 567–578.

Jordan, N. C., Levine, S. C., & Huttenlocher, J. (1995). Calculation abilities in young children with different patterns of cognitive functioning. *Journal of Learning Disabilities*, *28*, 53–64.

Jorm, A. F., Share, D. L., Matthews, M., & Matthews, R. (1986). Cognitive factors at school entry predictive of specific reading retardation and general reading backwardness: A research note. *Journal of Child Psychology and Psychiatry, 27*, 45–54.

Kavale, K. A. (1988). Learning disability and cultural disadvantage: The case for a relationship. *Learning Disability Quarterly, 11*, 195–210.

Kavale, K. A. (1995). Setting the record straight on learning disability and low achievement: The tortuous path of ideology. *Learning Disabilities Research & Practice*, *10*, 145–152.

Kavale, K. A., & Forness, S. R. (1994). Learning disabilities and intelligence: An uneasy alliance. In T. E. Scruggs & M. M. Mastropieri (Eds.), *Advances in learning and behavioral disabilities* (pp. 1–63). Greenwich, CT: Jai Press.

Kavale, K. A., & Forness, S. R. (2000). What definitions of learning disability say and don't say: A critical analysis. *Journal of Learning Disabilities*, *33*, 239–256.

Kavale, K. A., Fuchs, D., & Scruggs, T. E. (1994). Setting the record straight on learning disability and low achievement: Implications for policymaking. *Learning Disabilities Research & Practice*, *9*, 70–77.

Kavale, K. A., & Reese, J. H. (1992). The character of learning disabilities: An Iowa profile. *Learning Disability Quarterly*, *15*, 74–94.

Kellam, S. G., Rebok, G. W., Mayer, L. S., Ialongo, N., & Kalodner, C. R. (1994). Depressive symptoms over first grade and their response to a developmental epidemiologically based preventive trial aimed at improving achievement. *Development and Psychopathology*, *6*, 463–481.

Kirk, S. A. (1962). *Educating exceptional children*. Boston: Houghton Mifflin.

LeFevre, J. (2000). Research on the development of academic skills: Introduction to the special issue on early literacy and early numeracy. *Canadian Journal of Experimental Psychology*, *54,* 57–60.

Lerner, J. (1989). Educational intervention in learning disabilities. *Journal of the American Academy of Child and Adolescent Psychiatry*, *28*, 326–331.

Lukatela, G., & Turvey, M. T. (1998). Reading in two alphabets. *American Psychologist*, *53*, 1057–1072.

Lyon, G. R. (1995). Toward a definition of dyslexia. *Annals of Dyslexia, 45*, 3–30.

Lyon, G. R. (1996). The future of children: Special education for students with disabilities. *Learning Disabilities*, *6*, 54–76.

Lyon, G. R., Fletcher, J. M., Shaywitz, S. E., Shaywitz, B. A., Wood, F. B., Schulte, A., et al. (2001). Rethinking learning disabilities. In C. E. Finn, Jr., R. A. J. Rotherham, & C. R. O'Hokanson, Jr. (Eds.), *Rethinking special education for a new century*. Washington, DC: Thomas B. Fordham Foundation and Progressive Policy Institute.

MacMillan, D. L., Siperstein, G. N., & Gresham, F. M. (1996). Mild mental retardation: A challenge to its viability as a diagnostic category. *Exceptional Children*, *62*, 356–371.

McFadden, G. T. (1990). *Determination of the subtypal composition of several samples of learning disabled children selected on the basis of WISC FSIQ IQ level: A neuropsychological, multivariate approach*. Unpublished doctoral dissertation, University of Windsor, Ontario, Canada.

McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology*, *74,* 240–260.

Mercer, C. D., Jordan, L., Allsop, D. H., & Mercer, A. R. (1996). Learning disabilities definitions and criteria used by state education departments. *Learning Disability Quarterly*, *19*, 217–232.

Morris, R., & Fletcher, J. M. (1988) Classification in neuropsychology: A theoretical framework and research paradigm. *Journal of Clinical and Experimental Neuropsychology*, *10*, 640–658.

Morris, R. D., Stuebing, K. K., Fletcher, J. M., Shaywitz, S. E., Lyon, G. R., Shankweiler, D. P., et al. (1998). Subtypes of reading disability: Variability around a phonological core. *Journal of Educational Psychology*, *90*, 347–373.

Morrison, S. R., & Siegel, L. S. (1991). Arithmetic disability: Theoretical considerations and empirical evidence for this subtype. In L. V. Feagans, E. J. Short, & L. J. Meltzer (Eds.), *Subtypes of learning disabilities: Theoretical perspectives and research.* (pp. 189–208). Hillsdale, NJ: Lawrence Erlbaum Assoc.

Nation, K., Adams, J. W., Bowyer-Crane, A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology*, *73*, 139–158.

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language*, *37*, 85–101.

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.

National Joint Committee on Learning Disabilities (1988). Letter to NJCLD member organizations.

Nunes, T., & Bryant, P. (1996). *Children doing mathematics.* Oxford, England: Blackwell.

Oakhill, J. (1993). Children's difficulties in reading comprehension. *Educational Psychology Review*, *5*, 1–15.

Oakhill, J. V., Cain, K., & Bryant, P. E. (in press). The dissociation of single-word and text comprehension: Evidence from component skills. *Language and Cognitive Processes*.

Oakhill, J. V., Yuill, N., & Parkin, A. (1986). On the nature of the difference between skilled and less-skilled comprehenders. *Journal of Research in Reading*, *9*, 80–91.

Olson, R. K., Forsberg, H., Gayan, J., & DeFries, J. C. (1999). A behavioral-genetic analysis of reading disabilities and component processes. In R. M. Klein & P. A. McMullen (Eds.), *Converging methods for understanding reading and dyslexia* (pp. 133–153). Cambridge, MA: MIT Press.

Onatsu-Arvilommi, T., & Nurmi, J. E. (2000). The role of task-avoidant and task-focused behaviors in the development of reading and mathematical skills during the first school year: A cross-lagged longitudinal study. *Journal of Educational Psychology*, *92*, 478–491.

Pennington, B. F. (1999). Dyslexia as a neurodevelopmental disorder. In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders* (pp. 307–330). Cambridge, MA: MIT Press.

Pennington, B. F., Gilger, J. W., Olson, R. K., & DeFries, J. C. (1992). External validity of age versus IQ discrepant definitions of reading disability: Lessons from a twin study. *Journal of Learning Disabilities*, *25*, 639–654.

Perfetti, C. A. (1985) *Reading ability*. New York: Oxford University Press.

Perfetti, C. A., Marron, M. A., & Folz, P. W. (1996). Sources of comprehension failure: Theoretical perspectives and case studies. In C. Cornoldi and J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 137–166). Mahwah, NJ: Lawrence Erlbaum Assoc.

Rasanen, P. P., & Ahonen, T. (1995). Arithmetic disabilities with and without reading difficulties: A comparison of arithmetic errors. *Developmental Neuropsychology*, *11,* 275–295.

Robinson, N. M., Zigler, E., & Gallagher, J. J. (2000). Two tails of the normal curve: Similarities and differences in the study of mental retardation and giftedness. *American Psychologist*, *55*, 1413–1424.

Rodgers, B. (1983). The identification and prevalence of specific reading retardation. *British Journal of Educational Psychology*, *53*, 369–373.

Rourke, B. P. (1993). Arithmetic disabilities specific and otherwise: A neuropsychological perspective. *Journal of Learning Disabilities*, *26*, 214–226.

Rutter, M. (1982). Syndromes attributed to "minimal brain dysfunction" in childhood. *American Journal of Psychiatry*, *139*, 21–33.

Rutter, M. (1989). Isle of Wight revisited: Twenty-five years of child psychiatric epidemiology. *Journal of the American Academy of Child and Adolescent Psychiatry*, *29*, 633–653.

Rutter, M., & Yule, W. (1975). The concept of specific reading retardation. *Journal of Child Psychology and Psychiatry*, *16*, 181–197.

Sattler, J. M. (1993). *Assessment of children's intelligence and special abilities.* New York: Allyn and Bacon.

Satz, P., & Fletcher, J. M. (1980). Minimal brain dysfunctions: An appraisal of research concepts and methods. In H. Rie & E. Rie (Eds.), *Handbook of minimal brain dysfunctions: A critical view* (pp. 669–715). New York: Wiley Interscience Series.

Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K., Fletcher, J., Brady, S., et al. (1999). Comprehension and decoding: Patterns of association in children with reading difficulties. *Scientific Studies of Reading, 3*, 69–94.

Share, D. L., McGee, R., & Silva, P. A. (1989). I.Q. and reading progress: A test of the capacity notion of I.Q. *Journal of the American Academic of Child and Adolescent Psychiatry*, *28*, 97–100.

Share, D. L., McGee, R., & Silva, P. A. (1991). The authors reply. *Journal of the American Academy of Child and Adolescent Psychiatry*, *30*, 697.

Shaywitz, B. A., Fletcher, J. M., & Shaywitz, S. E. (1997). Attention-Deficit/Hyperactivity Disorder. In K. F. Swaiman & S. Ashwal (Eds.*), Pediatric neurology: Principles & practice* (pp. 585–597). St. Louis: C.V. Mosby.

Shaywitz, B. A., Holford, T. R., Holahan, J. M., Fletcher, J. M., Stuebing, K. K., Francis, D. J., et al. (1995). A Matthew effect for IQ but not for reading: Results from a longitudinal study. *Reading Research Quarterly*, *30*, 894–906.

Shaywitz, S. E. (1996). Dyslexia. *Scientific American*, *275*, 98–104.

Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., & Fletcher, J. M., & Makuch, R. (1992). Distribution and temporal stability of dyslexia in an epidemiological sample of 414 children followed longitudinally. *New England Journal of Medicine*, *326*, 145–150.

Shaywitz, S. E., Pugh, K. R., Jenner, A. R., Fulbright, R. K., Fletcher, J. M., Gore, J. C., et al. (2000). The neurobiology of reading and reading disability (dyslexia). In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. III; pp. 229–249). Mahwah, NJ: Lawrence Erlbaum Assoc.

Siegel, L. S. (1992). Dyslexic vs. poor readers: Is there a difference? *Journal of Learning Disabilities*, *25*, 618–629.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–293). Hillsdale, NJ: Lawrence Erlbaum Assoc.

Silva, P. A., McGee, R., & Williams, S. (1985). Some characteristics of nine-year-old boys with general reading backwardness or specific reading retardation. *Journal of Child Psychology and Psychiatry*, *20*, 407–421.

Skinner, H. A. (1981). Toward the integration of classification theory and methods. *Journal of Abnormal Psychology*, *90*, 68–87.

Slavin, R. E., Karweit, N. L., & Madden, N. A. (1989*). Effective programs for students at risk.* Boston: Allyn & Bacon.

Snow, C., Burns, M. S., & Griffin, P. (Eds.) (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Stanovich, K. E. (1991). Discrepancy definitions of reading disability: Has intelligence led us astray? *Reading Research Quarterly*, *26*, 1–29.

Stanovich, K. E. (1993). The construct validity of discrepancy definitions of reading disability. In G. R. Lyon, D. B. Gray, J. F. Kavanagh, & N. A. Krasnegor (Eds.), *Better understanding learning disabilities* (pp. 273–308). Baltimore: Paul H. Brooks.

Sternberg, R. J. (1991). Are we reading too much into reading comprehension tests? *Journal of Reading*, *34*, 540–545.

Stevenson, J. (1988). Which aspects of reading disability show a "hump" in their distribution? *Applied Cognitive Psychology*, *2*, 77–85.

Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children: The role of language comprehension and working memory skills. *Reading and Writing*, *4*, 245–256.

Stothard, S. E., & Hulme, C. (1996). A comparison of reading comprehension and decoding difficulties in children. In C. Cornoldi and J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 93–112). Mahwah, NJ: Lawrence Erlbaum Assoc.

Stuebing, K. K., Fletcher, J. M., LeDoux, J. M., Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (in press). Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal.*

Swanson, H. L. (1999). Reading research for students with LD: A meta-analysis of intervention outcomes. *Journal of Learning Disabilities, 32*, 504–532.

Swanson, H. L., & Siegel, L. (in press). Learning disabilities as a working memory deficit. *Issues in Education*.

Tager-Flusberg, H., & Cooper, J. (1999). Present and future possibilities for defining a phenotype for specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 1275–1278.

Tannock, R., Martinussen, R., & Frijters, J. (2000). Naming speed performance and stimulant effects indicate effortful, semantic processing deficits in attention-deficit/hyperactivity disorder. *Journal of Abnormal Child Psychology*, *28*, 237–252.

Taylor, H. G., & Fletcher, J. M. (1983) Biological foundations of specific developmental disorders:

Methods, findings, and future directions. *Journal of Child Clinical Psychology*, *12*, 46–65.

Tomblin, J. B., & Zhang, X. (1999). Language patterns and etiology in children with specific language impairment. In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders* (pp. 361–382). Cambridge, MA: MIT Press.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, J., et al. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology*, *91*, 579–594.

U.S. Department of Education (1999). 34 CFR Parts 300 and 303, Assistance to the states for the education of children with disabilities and the early intervention program for infants and toddlers with disabilities; final regulations. *Federal Register, 64*(48), 12406–12672.

U.S. Office of Education (1968). *First annual report of the National Advisory Committee on Handicapped Children*. Washington, DC: U.S. Department of Health, Education, and Welfare.

U.S. Office of Education (1977). Assistance to states for education for handicapped children: Procedures for evaluating specific learning disabilities. *Federal Register*, *42*, G1082-G1085.

van der Wissell, A., & Zegers, F. E. (1985). Reading retardation revisited. *British Journal of Developmental Psychology*, *3*, 3–9.

Vellutino, F. R. (1979). *Dyslexia: Theory and research*. Cambridge, MA: MIT Press.

Vellutino, F. R., Scanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition for reading disability. *Journal of Learning Disabilities*, *33*, 223–238.

Waber, D. P., Wolff, P. H., Forbes, P. W., & Weiler, M. D. (in press). Rapid automatized naming in children referred for evaluation of heterogeneous learning problems: How specific are naming speed deficits to reading disability? *Child Neuropsychology*.

Wadsworth, S. J., Olson, R. K., Pennington, B. F., & DeFries, J. C. (2000). Differential genetic etiology of reading disability as a function of IQ. *Journal of Learning Disabilities*, *33*, 192–199.

White, J. L., Moffitt, T. E., & Silva, P. A. (1992). Neuropsychological and socio-emotional correlates of specific-arithmetic disability. *Archives of Clinical Neuropsychology*, *7*, 1–16.

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children–Revised*. San Antonio, TX: Psychological Corp.

Whitehurst, G., & Lonigan, C. (1998). Child development and emergent literacy. *Child Development*, *69*, 848–872.

Whitehurst, G. J., & Massetti, G. (in press). How well does Head Start prepare children to learn to read. In E. Zigler & S. J. Styfco (Eds.), *The Head Start Debates (Friendly and Otherwise)*. New Haven: Yale University Press.

Wise, B. W., Ring, J., & Olson, R. K. (1999). Training phonological awareness with and without attention to articulation. *Journal of Experimental Child Psychology*, *72*, 271–304.

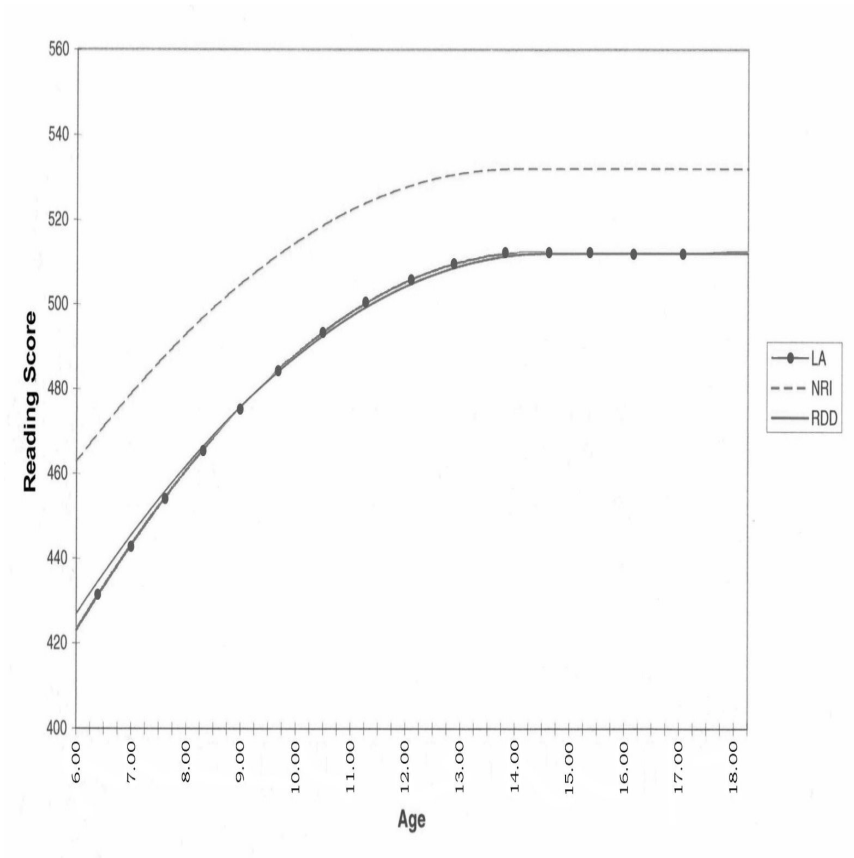
Wolf, M., & Bowers, P. G. (1999). The double deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, *91*, 415–438.

Wolf, M., Bowers, P. G., & Biddle, K. (2001). Naming-speed processes, timing, and reading: A conceptual review. *Journal of Learning Disabilities, 33*, 387–407.

Wood, F. B., Flowers, L., Buchsbaum, M., & Tallal, P. (1991). Investigation of abnormal left temporal functioning in dyslexia through rCBF, auditory evoked potentials, and positron emission tomography. *Reading and Writing: An Interdisciplinary Journal*, *3*, 379–393.

Woodcock, R. W., & Johnson, M. B. (1979). *Woodcock-Johnson Psycho-Educational Test Battery*. Boston: Teaching Resources.

Ysseldyke, J. E., Algozzine, B., Shinn, M. R., & McGue, M. (1982). Similarities and differences between low achievers and students classified learning disabled. *The Journal of Special Education*, *16*, 73–85.



Reading Score
560
540
520
500
480
460
440
420
400
6.00
7.00
8.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00
16.00
17.00
18.00
Age
LA
NRI
RDD

Figure 1

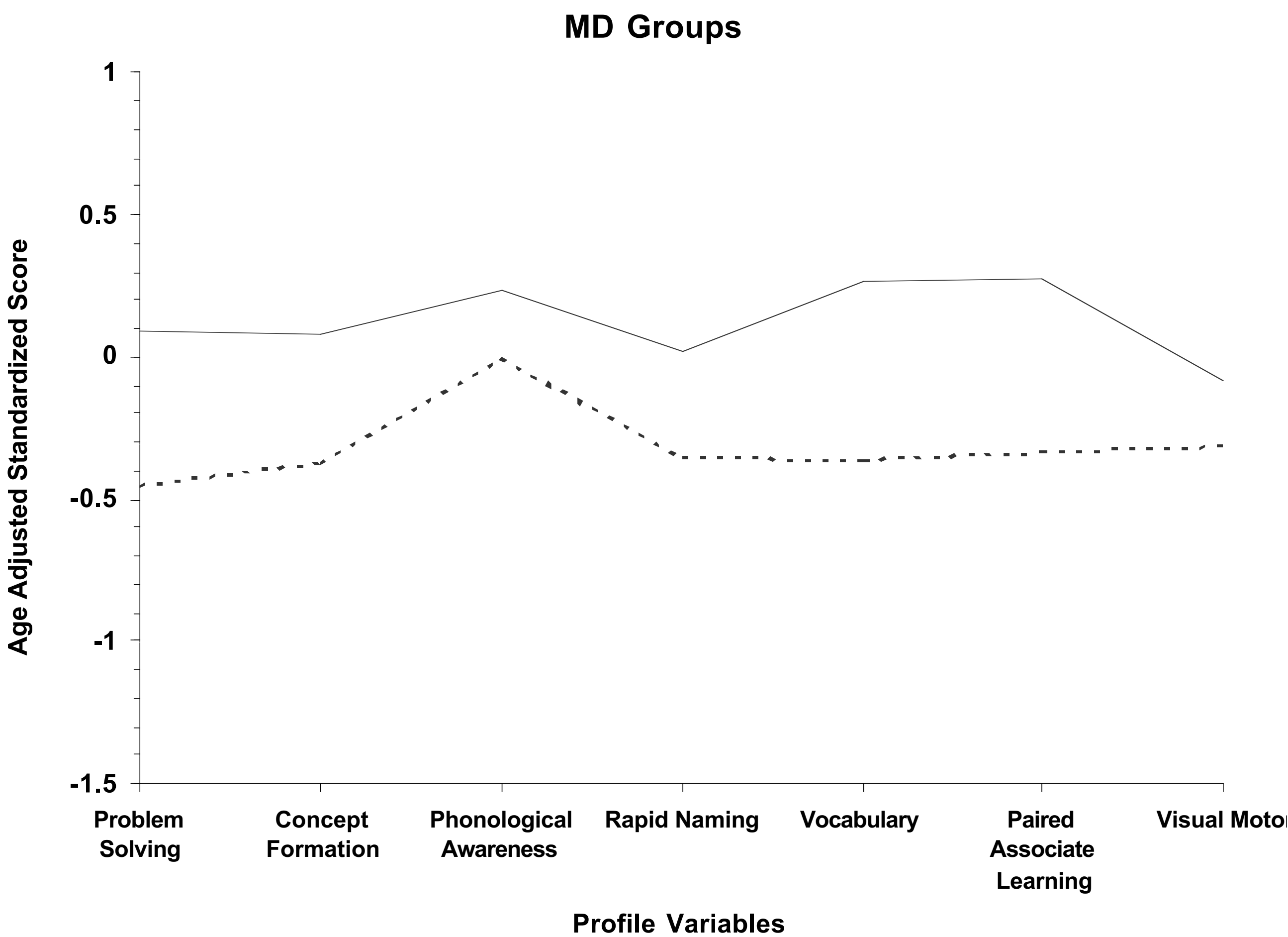

MD Groups
1
0.5
0
-0.5
-1
-1.5
Age Adjusted Standardized Score
Problem Solving
Concept Formation
Phonological Awareness
Rapid Naming
Vocabulary
Paired Associate Learning
Visual Motor
Profile Variables

Figure 2

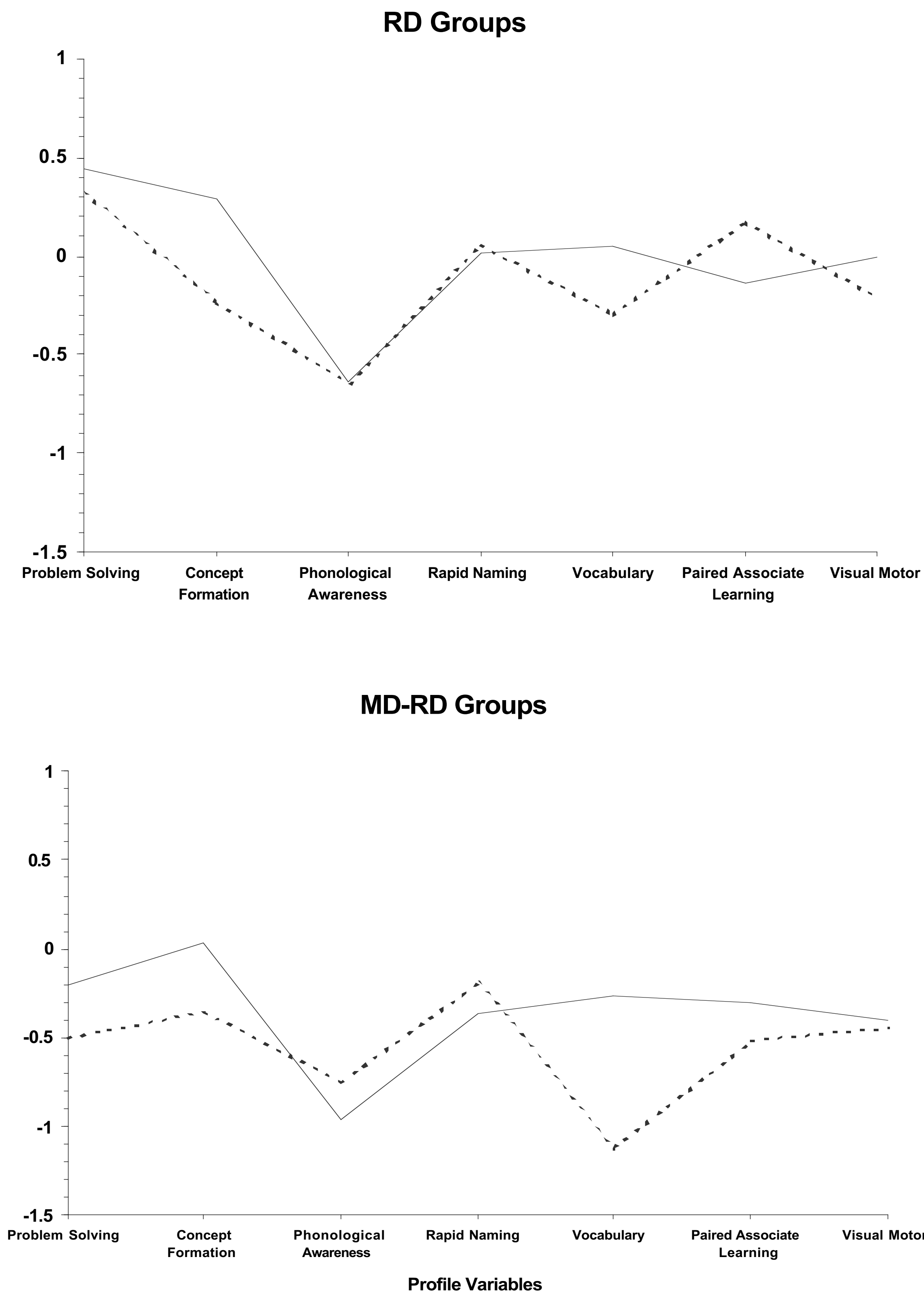

RD Groups
1
0.5
0
-0.5
-1
-1.5
Problem Solving
Concept Formation
Phonological Awareness
Rapid Naming
Vocabulary
Paired Associate Learning
Visual Motor
MD-RD Groups
1
0.5
0
-0.5
-1
-1.5
Problem Solving
Concept Formation
Phonological Awareness
Rapid Naming
Vocabulary
Paired Associate Learning
Visual Motor
Profile Variables

Figure 3

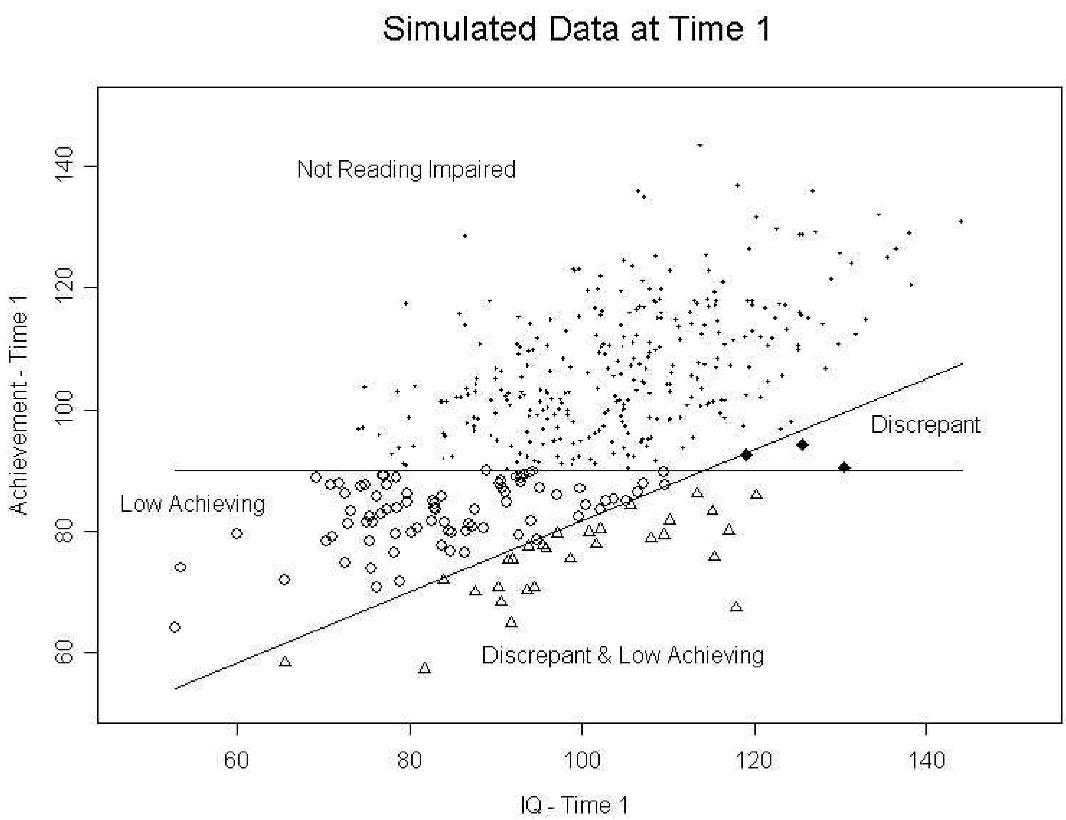

Simulated Data at Time 1
Not Reading Impaired
Low Achieving
Discrepant
Discrepant & Low Achieving
Achievement - Time 1
IQ - Time 1
60
80
100
120
140
60
80
100
120
140

Figure 4

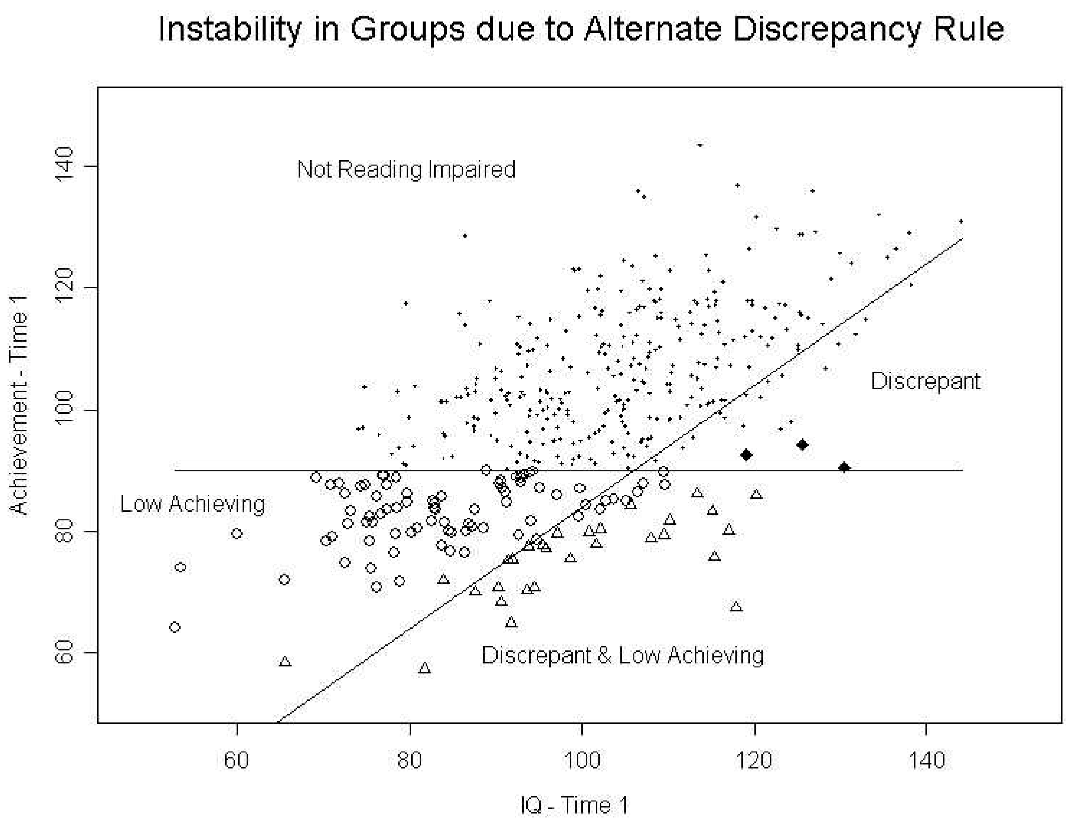

Instability in Groups due to Alternate Discrepancy Rule
Not Reading Impaired
Discrepant
Low Achieving
Discrepant & Low Achieving
Achievement - Time 1
IQ - Time 1
60
80
100
120
140
60
80
100
120
140

Figure 5

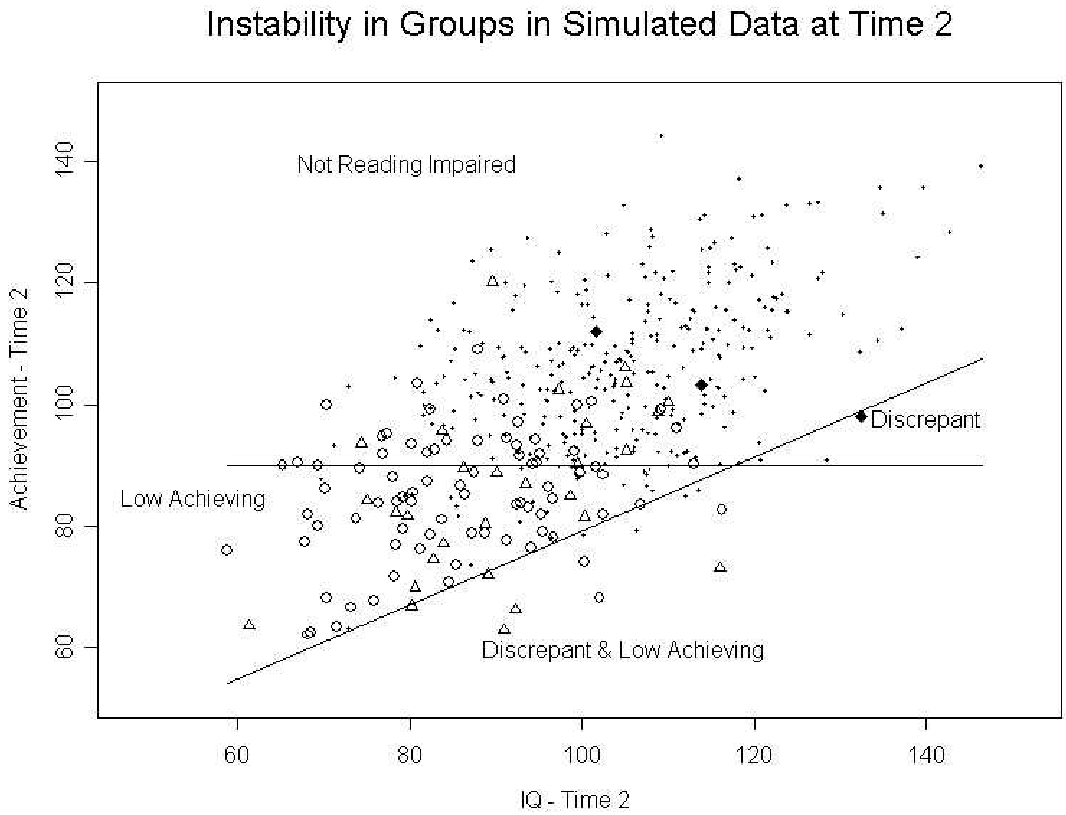

Instability in Groups in Simulated Data at Time 2
Not Reading Impaired
Discrepant
Low Achieving
Discrepant & Low Achieving
Achievement - Time 2
IQ - Time 2
60
80
100
120
140
60
80
100
120
140

Figure 6

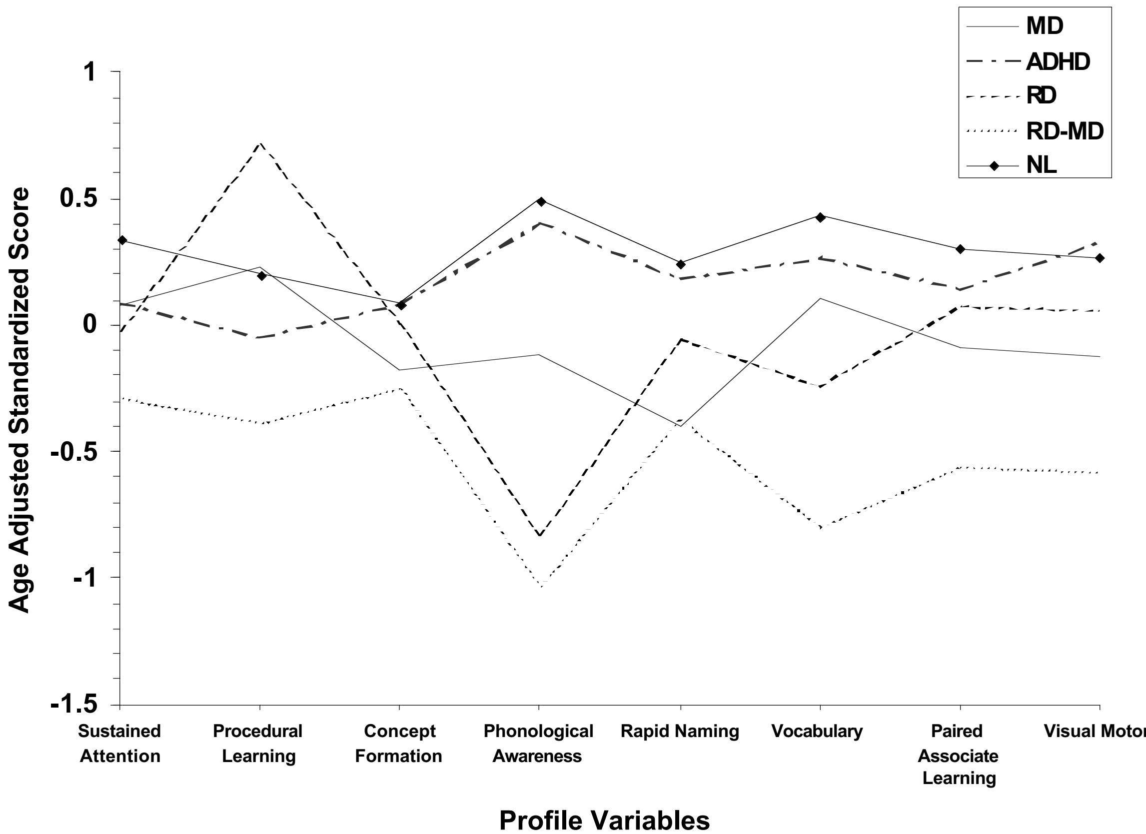

MD
ADHD
RD
RD-MD
NL
1
0.5
0
-0.5
-1
-1.5
Age Adjusted Standardized Score
Sustained Attention
Procedural Learning
Concept Formation
Phonological Awareness
Rapid Naming
Vocabulary
Paired Associate Learning
Visual Motor
Profile Variables

Figure 7

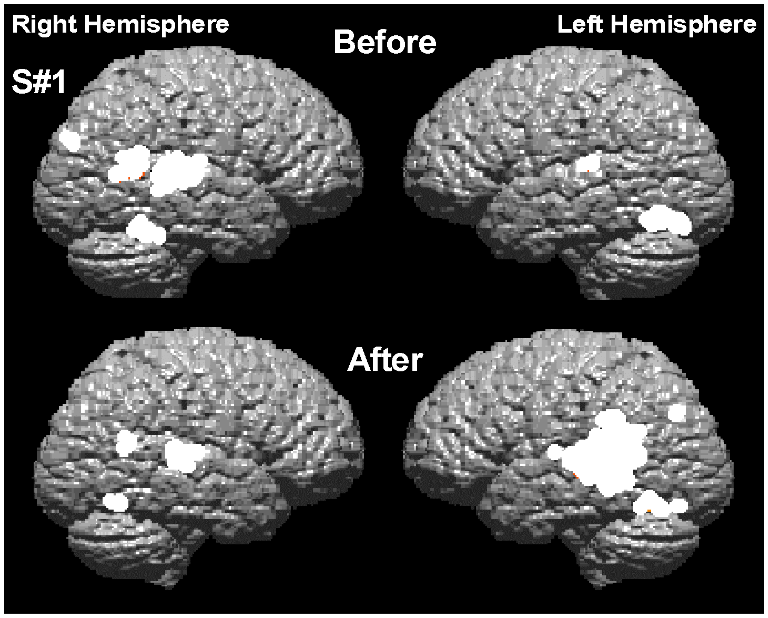


Figure 8

	Brain Pathology	No Known Brain Pathology	
	Average IQ Hydrocephalus	Poor Decoders–Poor Comprehenders	Good Decoders–Poor Comprehenders
Word recognition	Intact	Poor	Intact
Vocabulary knowledge	Intact	Poor	Intact
Reading comprehension	Poorer than word recognition	Poor	Poor
Literal recall of text	Poor	Intact	Intact
Inferencing	Poor	Poor	Poor
Primary source of inferencing failure	Accessing text- & knowledge-based information	Integrating text- & knowledge-based information	Integrating text- & knowledge-based information

Table 2