Classifications and Definitions for the Identification of Learning Disabilities: An Evaluation of the Research

Jack M. Fletcher, Ph.D.  
Department of Psychology  
University of Houston

RTI² Initiative 2013

jackfletcher@uh.edu
Disclosures

1. Author of Texas Primary Reading Inventory (Paul F. Brookes)
2. Author of Learning Disabilities: From Identification to Intervention
3. Research supported by NICHD grant, P50 HD052117, Texas Center for Learning Disabilities
4. Presentation not intentionally aligned with Tennessee standards. I am a neuropsychologist and scientist
5. Father of two grown (?) children
Objectives

1. Understand research underlying different methods for LD identification

2. Gain greater appreciation of the historical and conceptual basis for LD identification

3. Enhance capacity for conducting comprehensive evaluations of LD
What are Learning Disabilities? (how do I know one when I see one?)

- All disabilities have biological and social realities that vary with “disorder” and “person”
- Learning disabilities are dimensional—variation on normal development
- Model is obesity or hypertension, not measles and mumps
- “Disability” is a two pronged determination
Learning Disabilities is a Hypothetical Construct

- Essential aspect of construct is “unexpected underachievement”
- Constructs do not exist independently of how they are measured; all measures are imperfect indicators of constructs (latent variables)
- Measurement depends on definition
- Definitions and identification criteria derive from classifications
- Classifications reflect conceptual models
How LD is Identified and Treated Depends on the Conceptual Model

- **Neurological**: “Disorder of constitutional origin”: special signs

- **Cognitive Discrepancy**:
  - IQ-achievement discrepancy: cognitive discrepancy
  - Processing strengths and weaknesses: cognitive discrepancy

- **Instructional Discrepancy**
  - Low achievement: age-based discrepancy
  - Instructional response: intractability
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).
Is Charlie Brown LD? 1968

View of LD

What are the signs of LD? Identify a static, neurobiological disorder in order to intervene

"INDICATIONS OF POSSIBLE LEARNING DISABILITY...HAS A SLOW RECALL OF FACTS...MAKES SPELLING ERRORS...HAS AN UNSTABLE PENCIL GRIP..."
Learning disabilities are real! Stands up across definitional variation (doesn’t help identify individuals)

Children and adults with different forms of LD can be reliably and validly differentiated from each other, typical achievers, and other disabilities on cognitive correlates, response to intervention, and neural correlates

What happens when we apply these criteria to different classifications?
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).
What’s Wrong With IQ- Discrepancy?

- IQ- discrepant and non- discrepant low achievers do not differ significantly in behavior, achievement, cognitive skills, response to instruction, and neurobiological correlates once definitional variability accounted (Siegel, 1992; Stuebing et al., 2002).

- IQ does not predict intervention response (Stuebing et al., 2009).

- No difference in brain activation profiles (Tanaka et al., 2011)

- Status methods for identification may not be reliable based on a single assessment or cutpoint (Macmann et al., 1985; 1989; 1997; Francis et al., 2005)
RD Groups

- Problem Solving
- Concept Formation
- Phonological Awareness
- Rapid Naming
- Vocabulary
- Paired Associate Learning
- Visual Motor

Age Adjusted Standardized Score

IQ-Discrepant
Low Achievement
Discrepancy: Francis et al. (1996)
# Intervention Studies Addressing the Discrepancy Hypothesis

Strong relation with word recognition outcomes?

<table>
<thead>
<tr>
<th>Study</th>
<th>IQ</th>
<th>IQ- Discrepancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Foorman et al., 1998</td>
<td>No</td>
<td>--</td>
</tr>
<tr>
<td>2. Hatcher &amp; Hulme, 1999</td>
<td>No</td>
<td>--</td>
</tr>
<tr>
<td>3. Torgesen et al., 2000</td>
<td>No</td>
<td>--</td>
</tr>
<tr>
<td>4. Torgesen et al., 2001</td>
<td>No</td>
<td>--</td>
</tr>
<tr>
<td>5. Vellutino et al., 2000</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>6. Wise et al., 1999</td>
<td>Yes*</td>
<td>--</td>
</tr>
</tbody>
</table>

*Only 1 of 3 outcome measures

Stuebing et al. meta-analysis (Exceptional Children, 2009): overall unique R2 of .01 across 22 studies
Neurobiological Studies

- No difference in brain activation profiles of higher and lower IQ children with RD in reading-related areas (fMRI: Tanaka et al., 2011; MEG: Simos et al., 2013)

- Small difference in heritability of reading problems in higher and lower IQ children with RD (reading more heritable in higher IQ children; Wadsworth et al., 2010)

- DSM5 abandoned the IQ-discrepancy criterion (May, 2013)
Low Achievement method does not address unexpectedness

- Designate a cut point on the achievement dimension
- Strengths: Strong validity, linked to intervention, easy to implement
- Weaknesses: Cut point, does not measure the underlying construct (can’t differentiate subgroups of poor readers when the cause is known to be related to emotional difficulty, economic disadvantage, and inadequate instruction)
- Necessary but not sufficient: *Status models based on cutpoints for dimensional disorders may never be reliable for individuals*
Alternative Views: The “Third Method”

- Evaluate strengths and weaknesses in cognitive processes for inadequate responders to determine best Tx (ATI framework)
- Multiple “research-based” methods based on cognitive and achievement batteries:
  - Ability-Achievement Consistency (Flanagan);
  - Concordance-Discordance (Hale);
  - Discrepancy/Consistency (Naglieri)
- Hanson et al. (2008): “Research-based methods” recommended for Oregon schools
- Hale et al. (2010) survey of LD professionals: PSW methods needed not just for diagnosis, but also for treatment; mandated by statute
Problems with PSW Approaches

- Statute does not mandate that cognitive skills be assessed—just their manifestations.
- Little research on how PSW methods actually work and are related to instruction.
- Predicated on a straw person view of RTI (no standalone RTI identification method, comprehensive evaluation always required).
- Psychometric issues with discrepancy scores of any kind are well known, especially the use of rigid cut points, profile interpretations, difference scores, etc.
Simulation of PSW Methods (Stuebing et al., SPR, 2012)

- Created data sets where status of child as LD or not known; asked how well 3 PSW methods captured latent data at multiple differences.

- For all 3, number of children identified as LD low (about 2-3% depending on size of discrepancy).

- Specificity was generally higher than .85 and NPV was uniformly above .90. Sensitivity varied from poor (.17) to excellent (.91) across conditions and PPV was usually very low and never better than moderate.

For “not LD,” highly accurate (high specificity and few false negatives), but low PPV.
Of 10,000 assessments:

- **CDM**: 1,558 identified as LD (8,436 as not LD); 25 correct, so 1,533 are false positives and get the wrong treatment
- **DCM**: 362 identified as LD (9,638 not LD); 89 correct, so 273 are false positives and get the wrong treatment
- **XBA**: 678 would be identified as LD (9,322 not LD); 353 correct, 325 are false positives and get the wrong treatment

- Misinterpretation of significance tests; need to account for the test correlations; preoccupation with Type I error at the cost of significant risk for Type II errors; arbitrary cut points for discrepancy and low achievement
Agreement on LD identification between the C/DM and XBA methods at different low achievement cut points (Miciak et al., 2013)

<table>
<thead>
<tr>
<th>Approach</th>
<th>C/DM &lt; 85</th>
<th>C/DM &lt; 90</th>
<th>XBA &lt; 85</th>
<th>XBA &lt; 90</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/DM &lt; 85</td>
<td>-</td>
<td>62.1</td>
<td>30.0</td>
<td>13.6</td>
</tr>
<tr>
<td>C/DM &lt; 90</td>
<td>0.63</td>
<td>-</td>
<td>20.0</td>
<td>20.5</td>
</tr>
<tr>
<td>XBA &lt; 85</td>
<td>0.31</td>
<td>0.11</td>
<td>-</td>
<td>23.4</td>
</tr>
<tr>
<td>XBA &lt; 90</td>
<td>-0.04</td>
<td>0.03</td>
<td>0.22</td>
<td>-</td>
</tr>
</tbody>
</table>

Below diagonal = kappa; above diagonal = percentage overlap (total identified by both approaches/ total identified).
Performance on external reading variables of groups that met and did not meet PSW LD identification criteria
Evaluate cognitive characteristics of inadequate responders to Grade 1 Tier 2 intervention

Evaluate whether different outcomes measures yield subgroups that vary in cognitive characteristics (decoding vs. fluency)

Determine if there is unique or qualitative variability in cognitive skills not attributed to level of reading ability that would necessitate cognitive ability assessment
Criteria for Inadequate Response

- Norm Referenced Assessments of untimed word reading (WJIII Basic Reading) and timed word reading fluency \( \leq 25^{th} \) %tile

- CBM measure of passage reading fluency \( \leq 20 \) wcpm based on DIBELS end Grade 1 benchmarks (Continuous Monitoring of Early Reading Skills; CMERS)
Resultant Groups

- Decoding/Fluency (n = 29)
- Fluency (n = 75)
- Responders (n = 85)
- Typicals (n = 69)
- Assessed phonological awareness (CTOPP), rapid naming (CTOPP), speed of processing (Underlining), listening comprehension (CELF), syntactic comprehension/working memory (CELF), vocabulary/verbal reasoning (KBIT Verbal), and nonverbal problem solving (KBIT Matrices)
Cognitive Profiles of Inadequate Responders
Reading Level Evaluation

- Stanovitch and Siegel (1994): if cognitive skills contribute unique variance independent of level of reading ability, contrasts for IQ-achievement vs. low achieving poor readers should yield statistically significant beta weights after entering reading levels.

- WJIII Basic Reading, TOWRE composite, and CMERS wcpm entered into regression model predicting each cognitive variable, with a single contrast (adequate versus inadequate responders).

- No contrasts yielded statistically significant beta weights (rapid naming, \( t(256) = 1.88, p < .07 \)).
What do cognitive assessments add?

- Processing subtypes weakly related to intervention outcomes; little evidence that knowledge of cognitive strengths and weaknesses facilitates intervention (Pashler et al., 2010)

- No additional information not found in achievement profiles

- Cognitive deficits DO NOT reliably indicate biological causation; LD is an interaction of biological and environmental factors

- IQ when there is an issue about intellectual disability, autism spectrum disorder, or other disorder where IQ is directly relevant
New Alternatives: Response to Instruction (Intervention)

- Universal screening and serial curriculum-based assessments of learning in relation to instruction
- As one criterion, student may be LD if they do not respond to instruction that works with most students (i.e., unexpected underachievement)
- May identify a unique subgroup of underachievers that reflects an underlying classification that can be validated (Al-Otaiba & Fuchs, 2002; Vellutino et al., 2003)
- School-wide change - not just enhanced pre-referral services
What is Response to Intervention?

RTI is not:

- Just a special education initiative
- Only for students with disabilities
- Only for beginning reading
- Only for non-Title I and non-ESL students
- A new way to identify students with SLD
- A way of reducing costs or eliminating special education or the LD category
- This year’s summer reform or a short-term implementation based on “RTI in a Box”
- A way to fix schools with weak core instruction
Response to Intervention is:

- A set of processes for coordinating high quality service delivery in schools
- A multi-tiered, layered instructional approach that prevents problems first, and then brings increasingly intense interventions to students who don’t respond
- Making instructional decisions based on data
- Integrating entitlement programs with general education
- Providing relevant data for SLD identification
- Primary goal: Improving academic and behavioral outcomes for all students by eliminating discrepancies between actual and expected performance
Misconceptions of RTI

- Goal of RTI is to identify students as LD (RTI is a service delivery framework and identification is a by product of the process)
- Inadequate instructional response equates to special education eligibility (Instructional response is just one criterion for LD)
- Evaluation procedures fundamentally different (a comprehensive evaluation is required and most components of evaluation/eligibility are universal)
Comprehensive evaluation is required no matter what method is employed

- Data gathering process that includes child observation and may or may not use standardized tests

- In the context of RTI, goal not only special education eligibility, but to understand why the child has not responded to instruction

- In the context of RTI, instructional response data is routinely obtained (must be added to other identification methods in IDEA)

- Exclusionary criteria require consideration of other factors and may involve additional evaluation for other disabilities and language proficiency

- I think norm referenced assessments of achievement and behavior rating scales for screening are very helpful
RTI Adds:

- Documentation of parental notification and right to request an evaluation at any time
- Specification of learning strategies used to accelerate progress
- Some states add additional criteria for number of interventions, duration, and fidelity
LD Summit: Hybrid Method (Triangle Approach) to Identification (Bradley et al., 2002)

1. Establish Low Achievement

2. Evaluate Response to Instruction (Is underachievement expected?)

3. Apply the Exclusions

What is the validity of this hypothetical classification? (Low achievement is necessary, but not sufficient).

- [www.air.org/ldsummit](http://www.air.org/ldsummit)
Validity of the hybrid method (Fletcher et al., SPR, 2011)
Inadequate Responders: Tier 3 (baseline cognitive characteristics)
Adolescents: Tier 2 Cognitive Attributes

![Graph showing cognitive attributes for different tiers and responders.](image-url)
Grade 1 Intervention (pseudoword task)

Simos et al (Neuropsychology, 2005) - after Grade 1 intervention in Mathes et al. (RRQ, 2005)
Baseline MEG Patterns for Adolescent Adequate and Inadequate Responders
Rezaie et al., 2011
Reliability of the Hybrid Method

- If approach is to take a single assessment and set a cut point, identification of individual students will still be inadequately reliable.

- Attributes of LD (low achievement, inadequate instructional response) are dimensional (continua).

- Difficult to assess people in relation to set cut point.

- May be improved if multiple criteria are used and confidence intervals.

- How many resources should be devoted to finding the right student? Treat, then test.
Simulation of Agreement (10,000 Cases)

- Consider WJIII Basic and TOWRE composite in Fletcher et al. (2011); $r = .88$ (.94 if corrected for unreliability). Set cut points at 25th %tile: agreement ($k$) = .76
  - If correlation = 1.0, $k = 1.0$
  - 50th %tile, $k = .77$; 10th%, $k = .71$
  - If actual reliability (<.90), $k = .76$
  - Adjust for normative differences (sample mean above normative mean for WJ and below on TOWRE, $k = .39$
  - Sample size of 257, $k = .27-.51$
Actual Agreement

- WJ-TOWRE: $k = 0.38$
- WJ-CBM benchmark: $k = 0.25$
- CBM benchmark-TOWRE: $k = 0.61$
- Dual Discrepancy: $k = 0.21$ with WJ, $0.58$ with CBM benchmark, $0.60$ with TOWRE
Coverage

- Consider 104 inadequate responders as pool to be detected. How many not detected by each indicator?
  - WJ: .72
  - TOWRE: .14
  - CBM benchmark: .30.
  - Dual Discrepancy: .11 (but increases pool to 134, adding 29 inadequate responders and 1 typical (i.e., higher achievers))
Multiple Criteria

- CBM benchmark alone identified 14 children with reading scores on TOWRE, WJ, and other tests well above the average range (false positives?); this number increased dramatically with dual discrepancy.

- TOWRE and CBM benchmark agreed on 90/104 children, excluding those only identified by CBM or the 30 added by dual discrepancy (about 5’ of assessment time).

- Think about a pool; use multiple assessments; prioritize Type II over Type I errors (i.e., set the cut point high).
Identification issues are universal across methods

- No qualitative markers of LD (dimensional disorder)
- Measurement error (why do we persist with rigid cut points?)
- Instructional response may be a continuum; no qualitative markers of inadequate responders
- Specific issues in RTI are more than cut points and don’t equate to the adequacy of the measurement of instructional response
- How does the field move to informed decision making using multiple criteria and stop relying on psychometric methods?
RTI is not a panacea for problems with identification and intervention

- Key issue is implementation, especially enhanced instruction- it’s a scaling problem
- Linking general and special education/other entitlement programs is hard
- Identifying inadequate responders- still a continuum with potential cut point issues
- Resources must be redeployed
- Need more research on core instruction in math and written expression and tier 2/3 in math
- Knowledge base on inadequate responders is weak

Research is Evolving!!
But:

- RTI provides an alternative to cognitive (or even older neurological) conceptualizations of LD
- Directly linked to instruction and enhanced outcomes
- Cognition is related to LD and there are prominent neurological and genetic factors, but this knowledge does not yet facilitate identification or intervention
- RTI makes LD a real construct. We can argue about how to measure LD, but underlying constructs are real and survive definitional variability
Who is LD?

- The student who does not respond to quality instruction: *hard to teach, not unable to learn*

- Low achievement and inadequate instructional response

- Often preventable with early intervention

- Heritable, but neural systems are malleable
Can We “Psychometrize” Individual Identifications of LD? Not a New Question!

“Even though the psychometric difficulties may never be completely resolved, classification systems should at least be based on a coherent psychology of helping...there is no shortage of children who experience problems...Assessments and classifications can be guided by principles of intervention design with expected errors of judgment and measurement partially moderated through a recursive {sequential} system of recursive and empirical practices... (Macmann et al., 1988, p. 146)

“The real dilemma may be that procedures no more technically adequate than {formula-based procedures} are in wide use today. One wonders if a technically adequate solution to the problem of LD identification exists” (Danielson & Bauer, 1978, p. 175)
All professionals must...

- Focus on assessment of academic skills and move students to intervention as soon as possible- look at progress and *academic* strengths and weaknesses
- Address comorbid disorders and other factors
- Become experts on intervention
- Evaluate progress
- Reserve comprehensive evaluations for clearly inadequate responders
- Don’t get hung up on process- focus on results