

Differential diagnosis of dysgraphia, dyslexia, and OWL LD: behavioral and neuroimaging evidence

Virginia W. Berninger¹ · Todd L. Richards² · Robert D. Abbott³

Published online: 21 April 2015
© Springer Science+Business Media Dordrecht 2015

Abstract In Study 1, children in grades 4–9 ($N = 88$, 29 females and 59 males) with persisting reading and/or writing disabilities, despite considerable prior specialized instruction in and out of school, were given an evidence-based comprehensive assessment battery at the university while parents completed questionnaires regarding past and current history of language learning and other difficulties. Profiles (patterns) of normed measures for different levels of oral and written language used to categorize participants into diagnostic groups for *dysgraphia* (impaired *subword* handwriting) ($n = 26$), *dyslexia* (impaired *word* spelling and reading) ($n = 38$), or *oral and written language learning disability OWL LD* (impaired oral and written *syntax* comprehension and expression) ($n = 13$) or control *oral and written language learners (OWLs) without* specific learning disabilities (*SLDs*) ($n = 11$) were consistent with reported history. Impairments in working memory components supporting language learning were also examined. In Study 2, right handed children from Study 1 who did not wear braces (controls, $n = 9$, dysgraphia, $n = 14$; dyslexia, $n = 17$, OWL LD, $n = 5$) completed an fMRI functional connectivity brain imaging study in which they performed a word-specific spelling judgment task, which is related to both word reading and spelling, and may be impaired in dysgraphia, dyslexia, and OWL LD for different reasons. fMRI functional connectivity from 4 seed points in brain locations involved in written word

Electronic supplementary material The online version of this article (doi:[10.1007/s11145-015-9565-0](https://doi.org/10.1007/s11145-015-9565-0)) contains supplementary material, which is available to authorized users.

✉ Virginia W. Berninger
vwb@u.washington.edu

¹ Learning Sciences and Human Development, University of Washington, 322 Miller, Box 353600, Seattle, WA 98195-3600, USA

² Radiology, University of Washington, Seattle, WA, USA

³ Quantitative Studies and Measurement, University of Washington, Seattle, WA, USA

processing to other brain regions also differentiated dysgraphia, dyslexia, and OWL LD; both specific regions to which connected and overall number of functional connections differed. Thus, results provide converging neurological and behavioral evidence, for dysgraphia, dyslexia, and OWL LD being different, diagnosable SLDs for persisting written language problems during middle childhood and early adolescence. Translation of the research findings into practice at policy and administrative levels and at local school levels is discussed.

Keywords Dysgraphia · Dyslexia · Oral and written language learning disability (OWL LD) · Brain connectivity · Word-specific spelling

Introduction

Consensus has not been reached within and across countries for definitions of specific learning disabilities (SLDs) in general or those affecting written language acquisition in particular. The lack of consensus poses challenges for both researchers and educational practitioners worldwide. For example, even though *specific* in SLDs was historically intended to mean the student struggles with reading or another academic skill, despite otherwise normal development, frequently students with diagnosable pervasive developmental disabilities outside the normal range across multiple developmental domains are described as also having learning disabilities (Batshaw, Roizen, & Lotrecchiano, 2013). Also, in the United States (US) students are qualified for special education services under a single umbrella category, learning disabilities, despite the plural suffix of SLDs implying that not all learning disabilities are the same and the growing concern that state eligibility criteria are not sensitive to identifying all SLDs interfering with oral and written language learning. Regrettably, neither the researchers nor policy makers and government regulators can agree on definition for the various SLDs. The widely used approaches to defining SLDs—either discrepancy between the full scale score on a cognitive test and an achievement test or response to intervention (RTI)—have not resolved the definitional issues. On the one hand, researchers have not identified a specific size of discrepancy that alone is diagnostic in and of itself without additional evidence about which achievement skills are impaired. On the other hand, if RTI is not linked to the nature of an SLD, then failure to respond to whatever intervention is provided fails to provide diagnostic information regarding why the child struggles and how the intervention should be modified to facilitate RTI.

Moreover, epidemiological studies conducted by the Mayo Clinic in the US of the incidence of SLDs affecting written language learning, with and without co-occurring ADHD and/or math disabilities, in otherwise typically developing students, have shown that, regardless of definitions used, about 20 % of school age children and youth in the United States probably have some kind of SLD that may interfere with school learning at some time in their education (Katusic et al., 2005; Katusic, Colligan, Barbaresi, Schaid, Jacobsen, 2001; Katusic, Colligan, Weaver, Barbaresi, 2009; Stoeckel et al., 2013; St. Sauver, Katusic, Barbaresi, Colligan, Jacobsen, 2001; Yoshimasu et al., 2011, 2012). Given the frequency with which

these SLDs occur in the school age population, the overall research goal of the current study, which is part of a larger interdisciplinary programmatic research program on diagnosis and treatment of SLDs during middle childhood and early adolescence, was to use both behavioral and brain data to evaluate potentially converging evidence for differentiating among three of the most prevalent SLDs in school age populations: dysgraphia, dyslexia, and oral and written language learning disability (OWL LD), also referred to as specific language impairment (SLI). Even though brain data are reported, the focus was on developmental struggles in learning specific oral and/or written language skills rather than on acquired disorders after learning a skill, for example, through brain injury, disease, or stroke.

Although considerable progress has been made through early intervention with explicit phonological awareness and decoding instruction to prevent SLDs in many students (e.g., Foorman, Arndt, & Crawford, 2011; Lyytinen et al., 2004), there are still students with persisting SLDs in the upper elementary and middle school grades (e.g., Christensen & Wauchope, 2009; Lovett, Barron, & Frijters, 2013; Samuelsson et al., 2008). Indeed, brain differences have been identified between students with mild phonological difficulties who respond to early intervention and those who have persisting, more complex SLDs (Leonard, Eckert, Given, Berninger, & Eden, 2006). Thus, the specific aim of Study 1 was to evaluate whether an evidence-based, conceptually grounded assessment model for differential diagnosis of three SLDs (dysgraphia, dyslexia, and OWL LD) could categorize the referrals for persisting reading and/or writing problems in grades 4–9, despite earlier intervention, in students who otherwise were developing normally. The related specific aim of Study 2 was to evaluate whether these three diagnostic groups would also differ in fMRI functional connectivity during the same word-specific spelling task, thus providing converging brain evidence for the construct validity for these three SLD diagnoses.

Defining and diagnosing SLDs

Fundamentally, differential diagnosis is based on defining both what something is and what it is not and how it differs from other similar constructs. A word of Greek origin, *developmental dysgraphia* is the condition (suffix—*ia*) of having impaired (*dys*) letter form production through hand (*graph*) (*base word*), which impairs handwriting that in turn may interfere with learning to spell. Also a word of Greek origin, *developmental dyslexia* is the condition (suffix—*ia*) of having impaired (*dys*) word (*lexi* from *lexicon* or *mental dictionary*). Although dyslexia is often thought to be a reading problem, the persisting problem is spelling (Lefly & Pennington, 1991; Schulte-Korne et al., 1998) and affected individuals may have relative strengths in listening comprehension and oral expressions skills at the syntax level despite their word-level problems in decoding during reading and encoding during spelling. Both developmental dysgraphia and developmental dyslexia are typically first evident when handwriting or word decoding and spelling are first taught explicitly in kindergarten or first grade.

However, considerable research has shown that some other reading and writing problems are associated with significant impairments in syntax-level skills and may

emerge during the preschool years (Ellis & Thal, 2008; Mashburn, & Myers, 2010), but continue during the school age years when affected individuals may have not only continuing oral language problems but also written language problems (Apel, & Apel, 2011; Pennington & Bishop, 2009; Scott, 2010, 2011; Silliman & Scott, 2009). Such developmental OWL LD/SLI may manifest first as late onset of talking (Paul, Murray, Clancy, & Andrews, 1997; Thal, Bates, Goodman, & Jahn-Samilo, 1997; Thal, & Katich, 1996) or delayed combining of words due to weaknesses in syntax and related morphology (Catts, Bridges, Little, & Tomblin, 2008; Connelly, Dockrell, & Barnett, 2012; Scarborough, 2005). Although oral language disabilities may respond to treatment in the preschool years, the affected children are likely to have ongoing listening comprehension, oral expression, reading comprehension, and/or written expression problems, often related to morpho-syntactic levels of language and sometimes word finding, during the school years, unless identified and treated (see Nelson, 2010). Thus, developmental history including first emergence of an SLD and educational history of assessment, instruction, and ongoing manifestations of an SLD may be as relevant as test scores in defining and differentiating SLDs.

In addition, the current study on defining and differentiating three SLDs drew on one research group's prior cross-sectional and longitudinal assessment studies and instructional intervention studies with students in the early and upper elementary grades (ages 6–11) and middle school grades (ages 12–14) to define dysgraphia, dyslexia, and OWL LD (for review, see Berninger & Richards, 2010). The assessment studies showed that some students had significant problems with handwriting, which in turn interfered with their spelling and/or composing fluency, but not with reading. Others had significant difficulty with word reading (decoding pronounceable nonwords and identifying real words on a list without context clues), which in turn might affect reading comprehension, and spelling until the word reading and/or spelling problems were remediated; but they did not have trouble with listening comprehension or oral expression except for momentary inattention difficulties related to co-occurring ADHD in some children. Yet others had significant difficulty with syntax level listening and reading comprehension and oral and written expression. The intervention studies showed that the nature of these difficulties, at the subword, word, or syntax levels, affected their patterns of RTI for the same interventions aimed at multiple levels of written language (Berninger, 2008). For example, those with dyslexia showed slower RTI for decoding instruction (pronounceable nonwords) whereas those with OWL LD showed slower RTI on real word reading instruction. Even though phonological awareness and decoding instruction enabled RTI for students with dyslexia, students with persisting dysgraphia did not show RTI unless systematic instruction in letter writing legibility and automaticity was included; and students with OWL LD did not show RTI unless systematic instruction in syntax level skills for language by ear, by mouth, by eye, and by hand was included (see Berninger & O'Malley May, 2011).

Conceptual frameworks for defining and diagnosing

Cascading levels of language

For reasons just discussed, the conceptual framework underlying the current research was based on the level of language of the hallmark, defining impairment of each of three SLDs: *subword* level handwriting, *word* level reading/spelling, or *syntax* level comprehension and expression. However, for two reasons, this conceptual framework is referred to as a *cascading levels of language framework* in which the hallmark, defining impairments of each SLD are based on the highest level of language that is impaired—from subword, to word, to syntax. First, individuals may or may not show co-occurring impairment in the written language skills involving the smaller units of language below the hallmark level of impairment, but they do not show hallmark impairment in the larger unit(s) of language beyond their primary impairment (e.g., those with dyslexia may or may not have co-occurring dysgraphia, but do not have co-occurring OWL LD/SLI). Second, the area of primary impairment may interfere with acquisition of skills at higher levels of language but once the primary impairment is remediated, the problems at the higher level of language resolve; for example, impaired handwriting may interfere with learning to spell words or rate at which words can be produced during composing; but for those with only dysgraphia, the impaired handwriting does not interfere with learning to read and treating their handwriting problems may lead to improved spelling and composing more quickly than in dyslexia or OWL LD. In sum, the key concept in the cascading levels of language framework is that identifying the hallmark impaired level of language is critical to understanding an individual's profile of learning (language achievement) skills and planning appropriate, specialized instruction that is necessary in order to remediate the SLD. Without clear diagnosis, the necessary instructional components may be missed. That is, evidence-based differential diagnosis is treatment-relevant.

Multiple working memory components supporting language learning

The conceptual model was also based on a multi-component working memory system that supports language learning. Baddeley's (2003) original model of working memory has evolved from only two storage mechanisms—phonological and visual spatial—to multiple word form storage and processing mechanisms—phonological, orthographic, and semantic (Crosson et al., 1999) and phonological, orthographic, and morphological word forms (Richards et al., 2006) and a syntax buffer for accumulating word forms (Berninger, Raskind, Richards, Abbott, & Stock, 2008). The phonological loop has been re-conceptualized as a cross-code integration mechanism that supports language learning (Baddeley, Gathercole, & Papagno, 1998). The rapid automatic naming (RAN) task, a measure of phonological loop, requires rapid integration of letter word names (via mouth) with letter form codes (via eyes), as also does learning to name written words and objects in the external environment. The rapid automatic alphabet letter writing

task, a measure of the orthographic loop, assesses cross-code integration of letter forms in memory with sequential finger movements, as is needed to learn to write single letters or letters in words by hand. Moreover, supervisory attention in working memory is no longer considered a single function, but rather a panel of functions for focusing (selective), switching (flexible), and sustaining (maintaining) attention (cf. Baddeley, 2003); and self-monitoring as assessed with the n-back task Richards et al., 2009b).

Thus, Chomsky's (1965) proposed language learning mechanism, in which syntax played a fundamental role, may be a multi-component working-memory architecture that supports language learning at other levels of language as well, with (a) word storage and processing units (phonological for heard and spoken words, orthographic for read and written words, and morphological for bases and affixes in heard, spoken, read, and written words), (b) syntax storage and processing units for accumulating ordered words across time, (c) two loops for cross-code integration of internal codes with output systems through mouth (phonological loop) and hand (orthographic loop), and (d) supervisory attention (e.g., focused, switching, and sustaining attention) and self-monitoring. This language learning architecture supports the cross-code mapping (phonological, orthographic, and morphological) at the subword and word levels (e.g., Bahr, Silliman, Berninger, & Dow, 2012) underlying learning to read and write English, a morphophonemic orthography (Venezky, 1970, 1999).

Reconceptualizing cognitive measures

Niedo, Abbott, and Berninger (2014) investigated application of a model that included a multi-component working memory system for language learning and verbal comprehension to predicting levels of typical reading and writing achievement. Verbal Comprehension, re-conceptualized as translation of cognitions into oral language, was included based on past research showing that this index or factor score is the best predictor on cognitive tests of academic achievement measures in referred and unreferred samples (Greenblatt, Mattis, & Trad, 1990; Vellutino, Scanlon, & Tanzman, 1991). *WISC IV Verbal Comprehension* assesses ability to explain orally in words, phrases, and sentences what words mean, how two words index similar concepts, and why certain actions are needed or certain things are the way they are in the world. Thus, students have to translate their cognitions into oral language and the test does not assess only cognition. However, this translation is not the full story. Beta weights from multiple regressions using measures of phonological, orthographic, and morphological word storage and processing, syntax storage and processing, phonological and orthographic loops, and supervisory attention (focused and switching) were used to model the multi-component working memory system supporting language learning. Results showed that (a) adding the set of weighted working memory components to Verbal Comprehension substantially increased the amount of variance predicted for the same reading and writing outcomes; and (b) the working memory components collectively explained more variance than Verbal Comprehension alone. Thus, in the current research we assessed patterns not only in learning profiles (oral and written language

achievement) but also in phenotype profiles (working memory components supporting language learning which may have brain/genetic bases).

Summary of conceptual model

In this model, the deep structure is not in the syntax of the language, but rather in the higher order executive function for translating across the domains of cognition and oral language; and the full language learning mechanism involves both a translator and a multi-component working memory system supporting language learning across cascading levels of language—from subword to word to syntax—for language by ear, language by mouth, language by eye, and language by hand. The current research was designed to evaluate whether this model can be applied to diagnosing dysgraphia, dyslexia, and OWL LD.

Four research questions and measurement approach

Thus we built upon prior research to address four research questions with implications for translation of interdisciplinary research into educationally relevant diagnostic practices.

Study 1

The first research question was if students in grades 4–9 (ages 9–15) with persistent problems with written language beyond the first three grades (ages 6–8) could be classified into one of these three SLDs defined on the basis of level of language of impairment—dysgraphia (subword), dyslexia (word), or oral and written language learning disability (OWL LD) (syntax)—as assessed with normed tests (learning profiles) and parent questionnaires and reported developmental and educational histories. The second research question was whether patterning of measures of working memory components supporting language learning (phenotype profiles) differed across the three diagnostic groups—dysgraphia, dyslexia, and OWL LD. The third research question was whether each of these three SLD diagnostic groups differed from each other and from the control typical oral and written language learners (OWLs) on mean scores on separate measures of specific oral or written language skills, working memory components, or cognitive-linguistic translation.

Study 2

The fourth research question was whether these three diagnostic groups might also differ in neurolinguistic profiles for fMRI functional connectivity from four seed points (brain locations) associated with processing and producing written words on a word-specific spelling task to other brain locations. Research has shown that all three diagnostic groups tend to be impaired on the word-specific spelling task (recognizing the correctly spelled word when foils sound like real words if

pronounced) at the behavioral level, but it is not known whether they may show a different pattern of brain activation on this task depending on the diagnosed SLD.

Pattern approach to measurement

For both Study 1 and Study 2, the approach to data analysis sought *patterns in the data* and was grounded in recognition that measured variables may *both group into categories and distribute along different dimensions within categories*. That is, diagnosis was not based on low performance on a single skill or dimension along which scores distribute (as in low achievement alone) nor a simple difference between distributed scores on two dimensions (as in discrepancy between cognitive and achievement measures), but rather was based on comparisons of patterns of relative positioning within and across multiple distributions of age- or grade-referenced behavioral measures and reported history of challenges in learning oral and written language, prior interventions, and current issues (Study 1). The validity of the diagnosis was also evaluated with brain patterns on a common word spelling task (Study 2).

Study 1

Method

Participants

Ascertainment and screening interview Flyers announcing an opportunity to participate in a study of oral and written language learning for students with and without SLDs affecting oral and written language learning were distributed to local schools in close proximity to the university where the research was conducted. Teachers shared the flyers with parents of students whose parents they thought might be interested in participating. The flyers were also posted at the university where the research was conducted. Interested parents contacted the first author for further information about the study; and if they were interested in participating, an initial interview was conducted over the phone using a scripted set of questions approved by the Institutional Review Board (IRB) of the university.

The purpose of the initial interview was to determine if factors other than SLD could probably account for the persisting oral and written language learning problems, for example, pervasive developmental disability, other neurogenetic disorders like fragile-X or neurofibromatosis or PKU, sensory disorders like significant hearing loss or visual impairment, motor disorders like cerebral palsy or muscular dystrophy, spinal cord or brain injuries, substance abuse, medical conditions like epilepsy or other seizure disorders, etc. ADHD was not an exclusion criterion because it is known to co-occur with dysgraphia, dyslexia, and OWL LD. Also, questions were asked about developmental history and prior and current educational issues related to the persisting struggle with some aspect of written language (reading familiar or unknown words, reading comprehension, handwriting, spelling, and/or composition). In some cases

parents also volunteered siblings without SLDs who shared home environments with those with SLDs to serve as controls in the study or others at the university who are interested in the research volunteered their children who do not struggle with oral or written language learning.

Comprehensive assessment If the phone screening showed that a child did not meet exclusion criteria and by report appeared to be typically developing other than having a struggle with specific oral and/or written language skills or was experiencing no such struggles and parent granted informed consent and child gave assent, as approved by the IRB for human participants in research, an appointment at the university was scheduled for comprehensive assessment. The comprehensive assessment took on average 4 h, with refreshment and movement breaks interspersed. While the participating student completed the test battery that included cognitive, oral language and aural language, reading, writing, and related processing measures given in a standard order across participants, the parent completed questionnaires about developmental, medical, family, and educational histories and rating scales about their child's motor, language, social emotional, attention and executive function development.

Assignment to diagnostic and control groups Following the scoring of the comprehensive assessment test measures by graduate research assistants supervised by the first author, the first author examined whether the pattern of the learning profile test scores converged with (a) one of the three SLD diagnostic groups based on past research and the conceptual model described earlier or the control group, and (b) the reported history (timing of first observation of learning difficulty), nature of learning difficulties during preschool, the early grades, and recent/current grades. Also noted was whether parent reported that the child had received ongoing extra instructional services at school in general or special education and/or extra help outside school; however, it was not possible to evaluate whether that instruction was tailored to the student's learning profile and instructional needs. If the test results and parent-completed questionnaire and rating scale information supported one of these diagnostic classifications, the child was assigned to one of the four groups in the current study. All students who qualified for the comprehensive assessment in the current study could be assigned to one of the three diagnostic groups or control group. The working memory components were not used to assign to groups but rather were analyzed after group assignments to evaluate the potential patterning of these measures as a function of specific diagnostic groups or control group assignment.

Sample characteristics Altogether 29 females and 59 males in grades 4–9 (ages 9–15, $M = 12$ years 3 months) completed the comprehensive assessment battery. *Racial identities* of these students, as reported by their parents, were representative of the region from which the participants were recruited and included White ($n = 69$), More than One Race ($n = 14$), Asian ($n = 3$), Native Hawaiian or Other Pacific Islander ($n = 1$), and Black or African American ($n = 1$). Their mothers'

levels of education included high school graduate (2.2 %), more than high school but less than college (3.3 %), college (41 %), more than college (48.9 %), and unknown (4.4 %). Their fathers' levels of education included high school graduate (4.4 %), more than high school but less than college (7.8 %), college (41.1 %), more than college (36.7 %), and unknown (7.8 %). With the exception of children who were adopted and biological parent family history was unknown, parents reported family history of reading and/or writing problems.

Measures

Cognitive measures The *Wechsler Intelligence Scale for Children, 4th Edition (WISC IV)* (Wechsler, 2003) *Similarities, Vocabulary, and Comprehension* subtests needed to obtain a Verbal Comprehension Index score (test–retest reliability .93 to .95) were administered. Raw scores were converted to scaled scores, which were combined to obtain a standard score ($M = 100$, $SD = 15$). *Woodcock Johnson Psychoeducational Battery, 3rd Edition (WJ III)* (Woodcock, McGrew, & Mather, 2001a) *Concept Formation* (test–retest reliability of .77), which assesses inductive reasoning, that is, ability to abstract concepts from examples of the concepts, and *WJ III Analysis-Synthesis* (test–retest reliability .83), which assesses deductive reasoning, that is, ability to apply a concept or a rule to solve a problem, were also given. Both yield a standard score ($M = 100$, $SD = 15$).

Aural and oral language measures *WJ III Oral Comprehension* (Woodcock, McGrew, & Mather, 2001b) *Oral Comprehension* (test–retest reliability .88), which is an aural cloze task, requires supplying a word orally during pause in unfolding oral text, and *Understanding Directions* (no reliability reported in test manual), which assesses ability to understand and follow the spoken directions, were also given. Both yield a standard score ($M = 100$, $SD = 15$). Also given was *Clinical Evaluation of Language Function 4th Edition CELF IV* (Semel, Wiig, & Secord, 2003) *Formulated Sentences* (test–retest reliability .62–.71), which requires for each of multiple items constructing an oral sentence from three provided words; it yields a scaled score ($M = 10$ and $SD = 3$).

Reading measures Also given were *Woodcock Psychoeducational Battery 3rd Edition* (Woodcock et al., 2001b) *Word Identification* (test–retest reliability .95), for which the child is asked to pronounce a list of written real words without context clues, and *WJ III Word Attack* (test–retest reliabilities .73 to .81), for which the child is asked to pronounce a list of written pseudowords; for both tests, the score is based on accuracy. The *Test of Word Reading Efficiency (TOWRE)* (Torgesen, Wagner, & Rashotte, 1999) *Sight Word Efficiency Test* (test–retest reliability is .91), which measures the child's accuracy in pronouncing printed real words in a list within a time limit of 45 s, and *Pseudoword Efficiency Test* (test–retest reliability .90), which requires a child to read a list of printed pronounceable non-words accurately within a 45 s time limit, were also given to assess rate of accurate oral reading of real or nonwords. Also given was *WJ III Psychoeducational Battery* (Woodcock et al.,

2001b) *Passage Comprehension* (test–retest reliability is .85), a reading comprehension analogue of the oral cloze task, for which the task is to supply orally a missing word in the blank that fits the accumulating context of the sentence and preceding text. For all five reading measures, the raw scores are converted to standard scores ($M = 100$, $SD = 15$).

Handwriting measures On the *Alphabet Writing Task*, an experimenter-designed test, children are asked to handwrite in manuscript (unjoined letters) the lower case letters of the alphabet from memory as quickly as possible in alphabetic order, but to make sure others can identify the letters. The raw score is the number of letters that are legible and in correct order during first 15 s. The raw score is converted to a z -score ($M = 0$, $SD = 1$), based on research norms for grade (inter-rater reliability .97). On the *Detailed Assessment of Speed of Handwriting (DASH) Best and Fast* (Barnett, Henderson, Scheib, Schulz, 2007) the task is to copy a sentence with all the letters of the alphabet under contrasting instructions: one’s best handwriting or one’s fast writing (interrater reliability .99). Students can choose to use their usual writing—manuscript (unconnected) or cursive (connected) or a combination. Note that even though the task is to copy letters in word and syntax context, the scaled score ($M = 10$, $SD = 3$), is based on legibility for single letters within the time limits. In the current study, two testers reviewed all the scored handwritten measures to reach consensus on scoring.

Spelling measures *WJ III* (Woodcock et al., 2001b) *Spell Sounds* (reliability coefficient of .76) was used to assess ability to spell pronounceable nonwords without semantic meaning, which requires application of alphabetic principle and orthotactic knowledge of permissible positions within words for certain single letters or letter groups (Treiman and Kessler 2014). Raw scores are transformed to standard scores ($M = 100$, $SD = 15$). Three subtests of the *Test of Orthographic Competence (TOC)* (Mather, Roberts, Hammill, & Allen, 2008) were given. For the *Letter-Choice* subtest (test–retest reliability .84–.88), the task is to choose a letter in a set of four provided letters to fill in the blank in a letter series to create a correctly spelled real word (word-specific spelling). For *Sight Spelling* (test–retest reliability .91–.70), the task is to add one or several letters of one’s own choosing (not from a set of four provided letters) to a partially spelled word with missing letters to create a word-specific spelling for a real word. For the *TOC Homophone Choice (ages 9–12) or Word Choice (ages 13–16)* (test–retest reliability .72–.75), the task is to identify a correct spelling for a specific word; even though there are different norms according to age of child, the scaled scores regardless of age of the child were analyzed to assess word-specific spelling. The score on all TOC subtests is a scaled score ($M = 10$, $SD = 3$). *Wechsler Individual Achievement Test, 3rd Edition (WIAT III) Spelling* (Pearson, 2009) (test retest reliability .92) was also given on which the task is to spell in writing dictated real words, pronounced alone, then in a sentence, and then alone. The score is a standard score ($M = 100$, $SD = 15$).

Sentence composing measures For *WIAT III Sentence Combining* (test–retest reliability .81) (Pearson, 2009), the task is to combine two provided sentences into one well written sentence that contains all the ideas in the two separate sentences. For the *WJ III Writing Fluency* (Woodcock et al., 2001b) (test–retest reliability .88), the task is to compose a written sentence for each set of three provided words, which are to be used without changing them in any way. There is a 7 min time limit. For both measures the score is a standard score ($M = 100$, $SD = 15$).

Working memory—phonological, orthographic, and morphological word storage and processing measures For *Comprehensive Test of Phonological Processing (CTOPP)* (Wagner, Torgesen, & Rashotte, 1999) *Nonword Repetition* (test–retest reliability .70), the task is to listen to an audio recording of nonwords, which are pronounced one at a time and then repeat exactly the heard oral nonword, which contains English sounds but has no meaning; the score is a measure of phonological word form storage and processing. For the *Test of Silent Word Reading Fluency (TOSWRF)* (test–retest reliability is .92) (Mather, Hammill, Allen, & Roberts, 2004), the task is to mark the word boundaries in a series of letters arranged in rows. The score is the number of correctly detected and marked word boundaries in 3 min, which is a measure of orthographic word form storage and processing. For both tasks, the raw scores are converted to standard scores for age ($M = 100$, $SD = 15$). For the experimenter-designed *Comes From*, which is a measure of morphological word form storage and processing, the task is to judge whether or not a word is derived from a base word (Nagy, Berninger, & Abbott, 2006; Nagy, Berninger, Abbott, Vaughan, & Vermeulen, 2003). Example items include the following: Does *corner* come from *corn*? Does *builder* come from *build*? In both cases the first word contains a common spelling (er), but only in the second example does it function as a morpheme (a suffix that transforms a verb or action word into a noun or person word). Raw scores are transformed to z-scores ($M = 0$, $SD = 1$) based on research norms for elementary and middle school grades.

Working memory loops for integrating internal and output codes measures For *Rapid Automatic Letter Naming RAN* (test–retest reliability .90) (Wolf & Denckla, 2005), which is a measure of phonological loop for cross-code integration in language learning (Baddeley et al., 1998) and has a genetic basis (Rubenstein, Raskind, Berninger, Matsushita, & Wijsman, 2014), the task is to name lower case printed letters arranged in rows. The total score is the time required to name all the letters in all the rows. It is converted to a standard score with a mean of 100 and SD of 15. The alphabet 15 rapid automatic letter writing (see handwriting measures), has been validated in behavioral (Berninger, 2009) and brain research (Richards, Berninger, & Fayol, 2009a) as an index of the orthographic loop for linking orthographic codes in the mind’s eye with the sequential finger movements.

Supervisory attention of working memory measures For *Delis Kaplan Executive Functions D-KEFS* (Delis, Kaplan, & Kramer, 2001) *Color Word Form Inhibition* (reliability ranges from .62 to .76), based on the classic Stroop task, the task is to

read orally a color word in black and then name the ink color for a written word in which the color of the ink conflicts with the color name of the word (e.g., the word red written in green ink). The difference in time for reading the words in black and naming the color of the ink that conflicts with the name of the color word is an index of focused attention (inhibition of irrelevant information). Raw scores are converted to scaled scores for age ($M = 10$, $SD = 3$). For *Rapid Automatic Switching* (RAS)—letters and numerals (test–retest reliability .90) (Wolf & Denckla, 2005), the task is to name alternating lower case printed letters and written numerals arranged in rows. The total score is the time required to name all the alternating letters and numerals in all the rows and provides a measure of switching attention. It is converted to a standard score ($M = 100$ and $SD = 15$).

Diagnostic procedures for group assignment

The interdisciplinary, differential diagnosis scheme in Silliman and Berninger (2011) was used to classify the 88 students who completed comprehensive assessment into one of the diagnostic groups or the control groups. The criterion noted is the upper limit for qualifying for a diagnostic group; most fell below that cut-off and often substantially below and on more than two measures. All had reported history of past and current struggle with the relevant language skills for the assigned diagnostic group.

- (a) *Dysgraphia*—below $-2/3$ SD (25th %tile) on two or more handwriting measures unless noted in the results for one handwriting measure and handwriting during written composing; no current or past reading problems, and a parent reported history of past and current handwriting difficulties, but no preschool oral language problems;
- (b) *Dyslexia*—two or more oral word reading or written spelling measures below $-2/3$ SD (25th %tile) or below population mean and at least 1 standard deviation (15 standard score points) below cognitive-linguistic translation (see note at end of procedures); no current syntax level listening comprehension or oral expression difficulties; and a parent reported history of past and current word decoding/real word reading (oral and/or silent) and/or spelling problems but no preschool oral language problems; or
- (c) *Oral and written language learning disability, OWL LD*—below $-2/3$ SD on two or more measures of syntax-level listening comprehension, oral expression, reading comprehension, and/or oral and/or written syntax construction; and parent reported preschool history of oral language problems.
- (d) *Typical language learning controls*—reading, writing, and oral language skills across levels of language at or above standard score of $-2/3$ SD ; and no reported past or current history of oral and/or written language learning problems.

Note that nearly 75 % of the participants assigned to the dyslexia group qualified whether we used the criteria of two word level reading and spelling measures below

–2/3 SD but no evidence of syntax level problems in aural/oral or read/written language, or whether we used the criteria in a multi-generational family dyslexia study (two word level reading and spelling measures below the population mean and at least one standard deviation (15 standard score points) below the Verbal Comprehension Index, an indicator of cognitive-language translation. Whereas the first approach has the advantage of identifying dyslexia in those with average translation ability (a higher order executive function), the second approach has the advantage of not missing those who are twice exceptional (cognitive scores in superior or better range that can mask word reading and spelling problems on normed tests despite a history of persisting word reading or spelling problems (Berninger & Abbott, 2013; van Viersen, Kroesbergen, Slot, de Bree, 2014).

In addition to patterns in the learning profile for current oral language, reading, and writing achievement, the first author examined the developmental and educational histories provided by the parent, the phone interviews with parents documenting ongoing struggles that persisted despite prior extra help in and/or out of school, and recorded observations of the child by the team member who performed the assessments. Except for the typical oral and written language learners (controls), the participants talked openly and poignantly about how hard school was for them because of their struggles with writing and/or reading. So there was reasonable evidence across test scores, parent-report, and self-report that the SLDs affecting written language were persisting during middle childhood and adolescence.

Data analyses

Descriptive pattern analyses of individual profiles First, the individual profiles that were examined for each participant are summarized in Supplementary Table 1 for dysgraphia, Supplementary Table 2 for dyslexia, Supplementary Table 3 for OWL LD, and Supplementary Table 4 for control OWLS. Readers can access the supplementary tables through the URL link in the Results Section. Because measures differ in scaling properties—normed standard score ($M = 100$, $SD = 15$) or scaled score ($M = 10$, $SD = 3$), or research-generated z -score ($M = 0$, $SD = 1$)—scores in individual profiles are compared in reference to the range within a distribution in which a score falls, which has been shown to be more reliable across test administration than a single test score: *average* (–2/3 SD to just below +2/3 SD), *low average* (–1.33 SD to just below –2/3 SD), *below average* (below –1.33 SD), *above average* (2/3 SD to just below 1.33 SD), and *superior* (at or above 1.33 SD).

Inferential statistical analyses of groups Second, a series of ANOVAs, which can handle uneven sample size across groups, was performed, to evaluate how the three SLD groups differ from each other and from controls. To control for multiple comparisons for testing theory-driven predictions based on prior research, a confidence interval was set for $p \leq .001$. Results for those measures on which main effects for the four groups were statistically significant for the levels of language and working memory measures are summarized in Table 1 and for the translation

Table 1 Means and standard deviations for measures on which group effect statistically significant

Measure	Dysgraphia <i>M (SD)</i>	Dyslexia <i>M (SD)</i>	OWL LD <i>M (SD)</i>	Controls <i>M (SD)</i>	Main effect <i>F (df)</i>
Handwriting (subword)					
Alphabet 15z	-1.68 (0.68)	-1.46 (0.77)	-1.70 (.72)	-0.50 (1.01)	6.70 (3, 83)
<i>DASH copy fast</i>	6.41 (2.90)	6.93 (3.32)	7.09 (3.96)	11.09 (1.87)	6.26 (3, 85)
Spelling and reading (word)					
<i>WJ III spell sounds</i>	103.78 (16.91)	91.55 (10.45)	78.73 (26.72)	105.27 (12.33)	9.21 (3, 85)
<i>WIAT III spelling</i>	100.27 (19.05)	85.00 (12.00)	84.82 (10.06)	108.64 (12.19)	11.95 (3, 84)
<i>WJ III word ID</i>	109.81 (10.95)	97.41 (9.45)	88.18 (13.75)	111.27 (8.08)	16.86 (3, 83)
<i>TOWRE sight</i>	111.69 (14.49)	96.23 (13.11)	85.40 (12.25)	116.45 (13.49)	17.04 (3, 84)
<i>WJ III word attack</i>	106.58 (11.84)	94.70 (8.34)	91.36 (11.04)	106.64 (17.78)	12.22 (3, 84)
<i>TOWRE phonemic</i>	107.92 (17.01)	86.5 (11.90)	85.70 (10.77)	114.27 (17.52)	21.14 (3, 84)
Oral and written language (syntax)					
<i>WJ III oral comp</i>	113.31 (10.57)	113.03 (8.36)	91.18 (12.94)	108.64 (11.80)	14.88 (3, 84)
<i>WJ III understand dir</i>	107.04 (17.61)	102.00 (12.14)	82.00 (11.71)	97.00 (32.56)	5.77 (3, 83)
<i>CELF IV form sent</i>	10.35 (2.62)	11.28 (2.83)	5.09 (2.36)	12.73 (1.62)	19.84 (3, 84)
<i>WJ III pass comp</i>	105.85 (11.33)	98.63 (9.45)	76.27(13.65)	106.27 (9.25)	22.28 (3, 84)
<i>WJ III writing fluency</i>	98.65 (13.02)	93.33 (11.39)	82.10 (16.47)	109.82 (8.21)	9.98 (3, 82)
Working memory components					
Orthographic coding					
<i>TOWSRF</i>	101.31 (13.93)	91.25 (9.05)	85.70 (10.77)	109.55 (9.85)	12.96 (3, 83)
Morphological coding					
<i>Comes from</i>	-0.03 (.77)	-0.26(.95)	-1.86(1.29)	0.22 (0.46)	11.71 (3, 79)
Phonological loop					
<i>Wolf and Denckla RAN</i>	108.31 (11.71)	97.88 (13.11)	103.18 (9.90)	113.82 (10.46)	6.97 (3, 84)

Table 1 continued

Measure	Dysgraphia <i>M (SD)</i>	Dyslexia <i>M (SD)</i>	OWL LD <i>M (SD)</i>	Controls <i>M (SD)</i>	Main effect <i>F (df)</i>
Orthographic loop See <i>alphabet 15z</i>					
Supervisory attention	10.46 (3.33)	8.30 (2.92)	6.60 (3.86)	11.30 (2.54)	6.17 (3, 80)
<i>D-KEFS inhibition</i>	110.31(9.95)	96.38 (11.61)	95.91 (9.79)	114.00 (13.91)	13.23 (3, 84)

Measures on which groups differed at $p \leq .001$ were used to evaluate the cascading levels of language model for defining dysgraphia, dyslexia, and OWL LD (see text). Also see text for full name of WJ III Word ID and syntax measures for oral and written language

Table 2 Cognitive measures on which ANOVA main effect for groups was statistically significant

Cognitive Tests	Dysgraphia <i>M (SD)</i>	Dyslexia <i>M (SD)</i>	OWL LD <i>M (SD)</i>	Controls <i>M (SD)</i>	Main effect <i>F (df)</i>
Verbal comprehension	109.56 (15.24)	112.85 (4.42)	85.18 (10.83)	107.60 (10.74)	14.14 (3, 84)
Concept formation	103.69 (18.46)	109.13 (13.28)	86.36 (13.89)	105.67 (12.12)	6.64 (3, 82)
Analysis-synthesis	110.15 (19.41)	113.03 (14.78)	90.09 (17.25)	111.67(14.98)	5.63 (3, 82)

First measure from *WISC IV*. Second and third measures from *WJ III* and significance set at $p \leq .001$ (see text)

(cognitive to language or cognitive to nonverbal) measures are summarized in Table 2. Then for each of those comparisons that were statistically significant, the mean size of the differences is reported organized by levels of language, working memory, and cognitive measures in Table 3). Overall, two kinds of controls were used. The first control was using measures with published test norms or research norms to compare each participant to age or grade peers in interpreting relative performance along a distribution of scores for measures relevant to differential diagnosis. This approach allows interpretation despite the age range of the participants because each participant is compared to others of the same age or grade. The second control was comparing each of the three SLD groups not only to each other but also to a group of typical language learners across levels of language, working memory components, and kinds of cognitive translation. Patterns of findings across these comparisons were analyzed for common as well as unique impairments between the different SLD groups and between each SLD group and control group to identify both shared and differentiating impairments for defining different SLDs.

Results

First and second research questions

The initial analyses addressed the first research question as to whether participants could be classified into the three diagnostic groups or control group based on level of language impairment. Once classified, subsequent analyses addressed the second research question as to whether, and if so, how the diagnostic groups differ in the patterning of the working memory component or cognitive translation measures. See Supplementary Tables 1, 2, 3, and 4.

Dysgraphia Altogether 22 of the 26 identified as having dysgraphia, met the dual criteria of at least two hallmark language impairments affecting handwriting at or below $-2/3$ S and past and current history of persisting handwriting problem, with no evidence of reading problems. Of the four who did not, Participants 5 and 22 were very below average on *DASH Copy Fast*, Participant 10 could not write any of the letters from memory, and Participant 11 was very low on the alphabet 15 task. All four were reported to have significant difficulty with handwriting for writing

Table 3 Mean differences on measures on which SLD groups differed from each other or controls

Group comparison	Measures	Mean difference
Dysgraphia versus dyslexia	Word-level	
	<i>WIAT III spelling</i>	15.27
	<i>WJ III word identification</i>	12.40
	<i>TOWRE sight word</i>	15.46
	Working-memory-phonological loop	
	<i>Wolf and Denckla RAN</i>	10.43
	Working memory—switching attention	
	<i>Wolf and Denckla RAS</i>	13.93
	Dysgraphia versus OWL LD	Cognitive
<i>WISC III Verbal Comprehension</i>		24.38
Syntax-level		
<i>WJ III oral comprehension</i>		22.13
<i>WJ III understanding directions</i>		25.04
<i>CELF IV formulated sentences</i>		5.26
<i>WJ III sentence writing fluency</i>		16.55
<i>WJ III passage comprehension</i>		29.58
Word-level		
<i>WJ III spell sounds</i>		25.05
<i>WJ III word identification</i>		21.63
<i>TOWRE sight</i>		26.29
<i>WJ III word attack</i>		15.22
<i>TOWRE phonemic</i>		22.22
Working memory—coding		
Orthographic <i>TOSWRF</i>		15.61
Morphological comes from		1.83
Working memory—switching attention		
<i>Wolf and Denckla RAS</i>		14.40
Dyslexia versus OWL LD		Cognitive
	<i>WIAT III verbal comprehension</i>	27.67
	<i>WJ III concept formation</i>	22.77
	<i>WJ III analysis synthesis</i>	22.94
	Syntax	
	<i>WJ III oral comprehension</i>	21.85
	<i>WJ III understanding directions</i>	20.00
<i>CELF IV sentence formulation</i>	6.19	
Dyslexia versus OWL LD	Working memory—coding	
	Morphological coding	1.60
Controls versus dysgraphia	Subword	
	Alphabet 15z	1.18
	<i>DASH copy sentence fast</i>	4.68
	Working memory—orthographic loop See alphabet 15	

Table 3 continued

Group comparison	Measures	Mean difference
Controls versus dyslexia	Subword	
	Alphabet 15z	0.96
	<i>DASH copy sentence fast</i>	4.16
	Word level	
	<i>WIAT III spelling</i>	23.64
	<i>WJ III word identification</i>	13.86
	<i>TOWRE sight</i>	20.22
	<i>WJ III word attack</i>	11.94
	<i>TOWRE phonemic</i>	27.72
	Syntax level	
	<i>WJ III writing fluency</i>	11.17
	Working memory—coding	
	Orthographic <i>TOSWRF</i>	18.30
	Working memory—loops	
	Phonological <i>Wolf and Denckla RAN</i>	15.94
	Orthographic see alphabet 15z	
	Working memory—switching attention	
	<i>Wolf and Denckla RAS</i>	17.62
	Controls versus OWL LD	Cognitive
<i>WISC III verbal comprehension</i>		22.42
<i>WJ III concept formation</i>		19.31
<i>WJ III analysis synthesis</i>		21.58
Syntax level		
<i>WJ III oral comprehension</i>		17.46
<i>CELF IV formulated sentences</i>		7.64
<i>WJ III writing fluency</i>		27.72
<i>WJ III passage comprehension</i>		30.00
Word level		
<i>WJ III spell sounds</i>		26.54
<i>WIAT III spelling</i>		23.82
<i>WJ III word identification</i>		23.09
<i>TOWRE sight word</i>		31.05
<i>WJ III word attack</i>		15.28
<i>TOWRE phonemic</i>		28.57
Working memory—coding		
Orthographic <i>TOSWRF</i>		23.85
Morphological comes from		2.08
Working memory—supervisory attention		
Focused attention <i>D-KEFS inhibition</i>	4.70	
Switching attention <i>Wolf and Denckla RAS</i>	18.09	

Reported only if groups differed at $p \leq .001$ (see text). First group always higher

assignments in the classroom, as confirmed by observation during comprehensive assessment of the poor quality of handwriting on compositions. For all but two of the 26, alphabet writing 15, which is not only a marker of subword letter language impairment but also of working memory orthographic loop impairment, was in the low average or below average range. Of the 26, 21 fell in the low average or below average range on *DASH Copy FAST* and exactly half (13) fell in the low average or below average range on *DASH Copy Best*. Other working memory component impairments besides orthographic loop included phonological word storage and processing ($n = 1$), orthographic word storage and processing ($n = 7$), morphological word storage and processing ($n = 4$), and focused/selective attention (D-KEFS Color Word Inhibition) ($n = 5$). (See Supplementary Table 1).

Dyslexia Of the 38 classified as having dyslexia, all but 10 met the dual criteria of having at least two word reading and/or spelling measures at or below $-2/3$ SD and past and current history of word reading and spelling problems. However, of the ten who did not, they met the criteria used in multigenerational family genetics studies in which the child qualifying family for participation had word reading and/or spelling skills below the population mean and at least a standard deviation below measured Verbal Comprehension Index (e.g., Raskind et al., 2005); and the students had continued to have persisting reading and spelling problems at school. Of the 28 who met the criteria with word reading and spelling in low average or below average range, 15 had co-occurring dysgraphia. Of those who met the criteria used in the family genetics study for students with superior or better verbal reasoning, 4 had co-occurring dysgraphia. Overall about half the children with dyslexia had co-occurring dysgraphia. None met criteria for OWL LD.

For those with dyslexia, impaired phonological ($n = 6$) and morphological ($n = 5$) word storage and processing and focused/selective attention ($n = 13$) occurred as had also been observed in dysgraphia. Overall, impaired orthographic coding ($n = 14$) and orthographic loop ($n = 30$) were the most frequent working memory component impairments common to both dysgraphia and dyslexia. However, two working memory impairments were observed in dyslexia but not dysgraphia: phonological loop (RAN) ($n = 12$) and switching/flexible attention (RAS) ($n = 13$), suggesting that that these working memory impairments may be unique hallmarks of dyslexia not shared with dysgraphia when dysgraphia occurs alone rather than co-occurring with dyslexia. Also of interest to both the first and second research question, considerable variability was observed as to which specific spelling and/or word reading skills or working memory components were impaired in individuals. (See Supplementary Table 2).

Owl Ld All 13 who were classified as OWL LD met the dual criteria of at least two syntax level skills at or below $-2/3$ SD and past and current history of oral comprehension and/or expression difficulties. In all cases at least one of the impairments involved aural or oral language syntax. However, many had impairments in written syntax, for example, on Sentence Combining, which is sensitive to written expression of thought at the syntax level and is currently a topic

of much developmental and instructional research on writing (e.g. Myhill 2008; Saddler & Graham, 2005). Of those with OWL LD, 2 had co-occurring dysgraphia, 2 had co-occurring dyslexia, 5 had co-occurring dysgraphia and dyslexia, and 4 had no co-occurring dyslexia or dysgraphia. The most frequent working memory component impairments were orthographic loop ($n = 11$) and morphology word storage and processing ($n = 9$). Working memory impairments were also observed in orthographic word storage and processing ($n = 7$), phonological word storage and processing ($n = 4$), focused/selective attention (color word inhibition) ($n = 4$), switching attention (RAS) ($n = 3$), and phonological loop (RAN) ($n = 1$). (See Supplementary Table 3).

OWLs The typical oral and written language learners tended to score in the average or above ranges on all the levels of language and working memory components. (See Supplementary Table 4.) They did not have history of struggles with oral or written language.

Third research question

To address the third research question, the three diagnostic groups were compared to the controls and each other for each measure on which there was a significant main effect. Table 1 summarizes the between subjects Analyses of Variance (ANOVA) results for each of the levels of language measures and working memory components on which the main effect for groups was significant at $p \leq .001$. These included alphabet 15 and *DASH Copy Fast* for *subword level*, two measures of dictated spelling (pseudowords and real words) and four measures of oral reading (accuracy and rate for real word and pseudowords) for the *word level*, and five measures (two aural listening comprehension, one oral expression, one reading comprehension, and one written expression) for the *syntax level*. The working memory components that showed main effects for groups included orthographic coding, morphological coding, phonological loop, orthographic loop, and supervisory attention (focused/selective attention and switching/flexible attention).

Table 2 summarizes the main effects for groups on the three cognitive measures at the $p \leq .001$. Group effects were observed for *WISC IV Verbal Comprehension* (cognitive to oral language translation), *WJ III Concept Formulation* (inductive reasoning—abstracting concepts), and *WJ III Analysis-Synthesis* (deductive reasoning—applying principles or rules).

Table 3 reports the comparisons between each of the groups with the other groups on the leveled language skills, working memory components supporting language learning, and cognitive translation. For those measures on each these comparisons two groups at a time that were statistically significant at $p \leq .001$, the mean size of the difference is reported with the first group in the comparison scoring higher. The *dysgraphia group outperformed the dyslexia group* on measures of dictated spelling of real words and accuracy and rate of oral reading of real words as well as phonological loop and switching attention. The *dysgraphia group outperformed the OWL LD group* on one cognitive measure (WISC IV Verbal

Comprehension), all measures of processing and/or producing oral language, reading comprehension and written composition, dictated spelling of real words and accuracy and rate of oral reading of real words and pseudowords, orthographic and morphological storage and processing of words, and switching attention. The *dyslexia group outperformed the OWL LD group* on all three cognitive measures, three oral language measures, and morphological storage and processing of words.

The *OWL control group outperformed the dysgraphia group* on orthographic loop (alphabet 15z) and *DASH Copy Fast*. The *OWL control group outperformed the dyslexia group* on the same measures as it outperformed the dysgraphia group, but also on the following measures: spelling dictated real words, accuracy and rate of oral reading of real and pseudowords, syntax level sentence composition fluency, orthographic coding, phonological loop, and switching attention. The *OWL control group outperformed the OWL LD group* on each of the cognitive measures, syntax level oral comprehension and formulated sentences, syntax level reading comprehension and written sentence construction, dictated spelling of pseudowords and real words, accuracy and rate of oral reading of real and pseudowords, measures of orthographic and morphological storage and processing of words, and focused and switching attention.

Discussion

Evidence for singular SLD or plural SLDs Overall, the individual profile analyses support the plural suffix on specific learning disabilities. The individual student analyses (first and second research questions) showed that the sample could be classified into three SLD groups based on level of language impairment and preschool and school history of past, current, and persisting learning difficulties, consistent with Silliman and Mody (2008). However, some variation among participants within a diagnostic group as to which specific written language and working memory measures were impaired was also observed in the examined individual profiles, consistent with evidence for genetic heterogeneity in SLDs (Raskind, Peters, Richards, Eckert, & Berninger, 2012).

Group analyses (third research question) showed even stronger evidence for dysgraphia, dyslexia, and OWL LD differing in level of language impairment and working memory phenotype impairment. The dysgraphia group differed from the other groups on subword level handwriting skills. The dyslexia group differed from the other groups in word level reading and spelling and differed from the OWL LD group on syntax level skills. These results are consistent with those of Catts, Aldof, Hogan, and Ellis (2005), who made the evidence-based case that dyslexia and specific language impairment (SLI), as OWL LD is also referred to as, are not the same. In addition, some impaired working memory components varied across diagnostic groups. For example, impairments in phonological loop and switching/flexible attention were frequently observed in dyslexia but never in dysgraphia alone. Impairments in morphological word coding were more frequent in OWL LD (9) compared to dyslexia (5) and dysgraphia (4). Although impairments in phonological word coding were observed in all groups, they were more frequent in dyslexia (6), next highest in OWL LD (4), and lowest in dysgraphia (1).

Impairments in orthographic coding also occurred with greater frequency in dyslexia (14) than in OWL LD (7) or dysgraphia (7) and often occurred in students with dyslexia who did not meet criteria for co-occurring dysgraphia.

Evidence for multiple overlapping, cascading levels of language impairment Overall the findings are consistent with cascading levels of language impairment—subword level impairment in dysgraphia alone, but word- level and sometimes subword-level in dyslexia, and syntax-level impairment and sometimes subword and/or word levels in OWL LD. Both for dyslexia and OWL LD, co-occurring SLDs were found involving lower levels of language. Co-occurring SLDs in the same student in a study of reading, writing, and arithmetic problems were also reported by Moll, Kunze, Neuhoff, Bruder, & Schulte-Körne (2014). However, not only level of language but also the language system defined by input or output systems needs to be taken into account in defining SLDs. The subword, word, and/or syntax levels may be impaired for language by ear (listening), by mouth (oral expression), by eye (reading), and/or by hand (writing). Dysgraphia involves mainly impaired language by hand, whereas dyslexia involves mainly impaired language by eye, and OWL LD involves mainly impaired language by ear and by mouth.

Commonalities across diverse SLDs Despite the variations among the three SLDs, they also share some common features (cf., Bishop, & Snowling, 2004). For example, the most common working memory impairment across all three SLD diagnostic groups was the orthographic loop. Of the 88 students in the sample with diagnosed SLDs, 65 showed evidence of impaired orthographic loop. The high co-occurrence of impaired orthographic loop across diagnostic groups and individuals may be related to this measure assessing *automaticity in letter formation* and *automaticity in accessing, retrieving, and producing ordered letter symbols*, both of which have been shown to play an important role in learning to spell and read words and as well as write letters (Berninger, 2009). As shown in the ground-breaking research by Schneider and Shrifin (Schneider & Shiffrin, 1977; Shrifin & Scheider, 1977) early in cognitive psychology, initial processing tends to be briefly on an automatic pilot, whereas subsequent processing engages more resource demanding controlled, strategic processing.

Also, across all three SLDs, orthographic storage and processing and morphological storage and processing impairments were more common than phonological storage and processing impairments in this sample of students with persisting reading and/or writing problems during middle childhood and early adolescence. This result may indicate that orthographic and morphological processing may be a relatively ignored source of difficulty in older students with persisting SLDs and one which requires specialized instruction during middle childhood and adolescence. Although progress has been made in educating educators about the importance of phonological awareness and phonics, work remains to educate them about the role of morphological and orthographic awareness in learning to read and spell a morphophonemic orthography. The alternations in English for applying phonics (grapheme-phoneme correspondence) in the early grades for learning to read orally

are often interpreted as evidence that English is an opaque, irregular orthography, but it is the case that not only English but most languages have a morphophonemic orthography (Cahill, Tiberius, Herring, 2013); and more than alphabetic principle is involved, especially for silent reading in the upper grades. Thus, a critical skill after the middle childhood transition to silent reading for both students with and without SLDs is learning the interrelationships among morphology, orthography, and phonology for the language used at school.

Evidence-based use of cognitive tests In the current study cognitive measures were administered, but SLDs were not defined on basis of difference between cognitive and reading or writing scores or solely on that basis in the case of verbally gifted with dyslexia (reading and spelling also had to be below the population mean). Rather, cognitive measures were used for three purposes: (a) to document that the student's cognitive skills were at least in the lower limits of the normal range, making it unlikely that the student had a developmental disability rather than an SLD; (b) to modify flexibly how far below average language skills had to be to be considered an impairment for verbally gifted students with dyslexia (Berninger & Abbott, 2013; van Viersse et al., 2014); and (c) to identify weaknesses in translation of cognition into oral language as may characterize some individuals with OWL LD. Although neither those with dysgraphia nor those with dyslexia differed from the controls on cognitive measures, those with OWL LD did differ from controls and differed on both what is generally regarded as verbal cognition (first measure in Table 2) and nonverbal cognition (last two measures in Table 2). That may be because even though the latter two measures require a nonverbal response, they also require processing heard language in the instructions for the various items.

It does not follow that the OWL LD group has cognitive deficiencies. Rather, they have impairments in their oral language skills for answering questions to express their cognitions on verbal measures; and for listening comprehension to understand the instructions and items administered with oral instructions even though oral responses are not required. That is, they are impaired in *cognitive to linguistic translation*. This impairment in the higher order executive function of translation may also affect other domains including nonverbal. The findings in Table 2 serve as a reminder that simple calculation of discrepancy between cognitive and academic achievement measures is not sufficient to identify specific kinds of SLDs. Students with OWL LD may not show discrepancy using that definitional approach. It does not follow, however, that students with OWL LD are less intelligent nor that what cognitive measures assess is irrelevant to academic achievement. Rather, low scores on these measures in students with OWL LD calls attention to their need for explicit instructional and learning activities to develop the oral language skills needed for a variety of educational purposes that depend on being able to translate one's cognitions into oral language or even nonverbal formats with verbal instructions.

Future translation research directions for clinical and educational practice Not all students in the US with dysgraphia, dyslexia, or OWL LD are being identified

and provided with appropriate instruction throughout schooling to ensure they receive appropriate instruction for both their oral language and written language instructional needs as curriculum requirements change across the grades (e.g., see Treiman & Kessler, 2014). The current practices for determining eligibility for special education do not take evidence-based differential diagnosis and treatment planning into account, but research shows that doing so can transform struggling students with OWL LD or dysgraphia or dyslexia into successful ones (e.g., Berninger & O'Malley May, 2011; Berninger et al., 2008). Moreover, not all reading problems are dyslexia, and preventing and treating persisting dyslexia requires spelling as well as reading instruction (e.g., Berninger, 2008, 2009). Given the heterogeneity underlying behavioral expression of SLDs across and within diagnostic groups, it is unlikely that a “one size fits all” approach to diagnostic procedures is evidence-based or can be reduced to algorithms that can be legally defined without professional judgment for the case at hand in the context of the profile for the whole child.

Study 2

Method

Participants

All participants in Study 1 were invited to participate in Study 2, a brain imaging study, if they were right handed and did not wear metal that could not be removed. Altogether 45 children (controls, $n = 9$, dysgraphia, $n = 14$; dyslexia, $n = 17$, and OWL LD, $n = 5$) met these requirements: their parents granted informed consent and they granted assent, approved by the Institutional Review Board; and they could tolerate the enclosed scanner and lay still to complete the word-specific spelling task during scanning.

fMRI imaging task

Task A word-specific spelling task was used as in prior fMRI BOLD studies with children with dyslexia (Richards et al., 2005) or dysgraphia (Richards et al., 2009a) of the same age as the children in the current study: decide if a letter string, which if pronounced sounds like a real word with meaning, is a correctly spelled real word. Foils sound like a real word but are not correctly spelled. This word-specific spelling task, which is used in much behavioral research on dyslexia beginning with the seminal work by Olson, Forsberg, Wise, and Rack (1994), takes into account not only orthographic, phonological, and morphological knowledge but also vocabulary meaning (e.g., Ehri, 1980), and may underlie oral word reading fluency and/or written spelling problems (Bowers & Wolf, 1993). For example, research has shown that students with dysgraphia without reading problems have difficulty with word-specific spelling during middle childhood/adolescent transition (Berninger & Hayes, 2012) as also do students with dyslexia (Berninger, Nielsen, Abbott, Wijsman, &

Raskind, 2008) and OWL LD (Berninger & O'Malley May, 2011). For dysgraphia, handwriting interferes with writing words in learning to spell. For dyslexia, hallmark impairments in phonological and orthographic coding can interfere with learning alphabet principle for word decoding in reading and word encoding in spelling. For OWL LD, impaired morphology or semantic/vocabulary knowledge may interfere with learning to spell.

Imaging procedures

fMRI functional connectivity and human connectome paradigm This study is part of a larger study grounded in the recent brain imaging paradigm shift to the human connectome (e.g., Caspers et al., 2012). Using state of the science techniques for fMRI functional connectivity (Eickhoff, Heim, Zilles, Amunts, 2006; Eickhoff et al., 2005, 2007; Toga, Thompson, Mori, Amunts, & Zilles, 2006), the research team analyzed acquired images for waves of connectivity from four seed points (specific brain locations) based on a metaanalysis of brain regions involved in written word processing and production: left occipital temporal, left supramarginal gyrus, left precuneus, and left inferior frontal gyrus (Purcell, Turkeltaub, Eden, Rapp, 2011). Left occipital temporal brain region has been shown in much brain research to be a word form region for initial processing of visible language. Left supramarginal gyrus in parietal lobe is involved in storing words and their parts in working memory. Left precuneus in parietal lobe is involved in reflecting upon the word parts and their interrelationships, that is, linguistic awareness. Left inferior frontal gyrus (Broca's area) in frontal lobe is involved in the executive functions for coordinating language systems throughout the brain. Of interest was whether the neurobiological profiles (patterns of fMRI functional connectivity based on number of regions to which there is a functional connection in time from a specific seed point and based on the specific region to which the seed point shows functional connectivity) differ across each diagnostic group (dysgraphia, dyslexia, OWL LD) with the control OWLs. Finding evidence that there are different patterns of fMRI functional connectivity among the three SLD diagnostic groups and control group would provide converging brain evidence for the behavioral evidence in Study 1.

Data acquisition T2*-weighted functional images (fMRI) (33 slices, 80×78 matrix, 240×240 FOV, 3 mm isotropic voxels) were acquired on a Philips Achieva 3T scanner with a 32-channel SENSE head coil using an EPI sequence (TR = 2 s, TE = 25 ms, flip angle = 79°). A 3D MPRAGE was collected for anatomical co-registration. During functional scans, participants were instructed to complete the word judgment task on which they had been trained prior to scanning. The task was presented with self-pacing for 2 min (60 time points).

Preprocessing Functional images were corrected for motion using FSL MCFLIRT software <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT>, then high-pass filtered at $\sigma = 20.83$. Motion scores (as given in the MCFLIRT report) were computed for each participant and average motion score (mean absolute displacement) for each of

the groups: Motion scores (as given in the MCFLIRT report) were computed for each participant and average motion score (mean absolute displacement during the reading fMRI) for each of the groups: control 1.18 ± 1.36 voxels, dysgraphic 1.17 ± 0.91 voxels, dyslexic 1.22 ± 0.89 voxels, OWL LD 1.24 ± 0.57 voxels). Based on these results, groups did not differ in motion. Spikes were identified and removed using the default parameters in AFNI's 3dDespike. Slice-timing correction was applied with FSL's slicetimer and spatial smoothing was performed using a 3D Gaussian kernel with $\text{FWHM} = 4.0$ mm. Time series motion parameters and the mean signal for eroded (1 mm in 3D) masks of the lateral ventricles and white matter (derived from running FreeSurfer's recon-all on the T1-weighted image) were analyzed. Co-registration of functional images to the T1 image was performed using boundary based registration based on a white matter segmentation of the T1 image through `epi_reg` in FSL. The MPRAGE structural scan was segmented using FreeSurfer software; white matter regressors were used to remove unwanted physiological components.

Individual functional connectivity maps were generated for each of the four different seed points in left visual cortex, left temporooccipital, left supramarginal, and left inferior frontal gyrus). fMRI time-series were averaged within regions of interest (ROI) formed from a 15 mm sphere centered at each seed. The averaged time-series at each ROI was correlated with every voxel throughout the brain to produce functional connectivity correlation maps, converted to z -statistics using the Fisher transformation. When a functional connectivity value is reported in Supplementary Table 5, it was computed from the correlation map/Fisher transformation.

A group fMRI connectivity map was generated for each of the 4 groups using a global design matrix as part of the GLM model in software randomise as describe by FSL guidelines for higher group level analysis as show by this weblink http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FEAT/UserGuide#Single-Group_Average_with_Additional_Covariate. This model generated the single-group average for each for the 4 groups (controls, dysgraphics dyslexics, and OWL LD) with a gender covariate.

Results

Number of regions to which functional connections observed from each seed point As shown in Supplementary Table 5, on the same word-specific spelling task, the four groups differed in number of connections from each of the four seed points to other brain regions. For the *left occipital temporal seed point*, those with OWL LD showed the least connectivity (to 21 regions); the controls somewhat more connectivity (to 26 regions); but those with dysgraphia (to 42 regions) showed considerably more connectivity; and those with dyslexia (to 79 regions) showed the most connectivity. For the *left supramarginal gyrus seed point*, those with OWL LD showed underconnectivity (to 14 regions) compared to the controls (to 46 regions); those with dysgraphia had even more connectivity (to 75 regions); and those with dyslexia had the most connectivity (to 110 regions). For the *left precuneus seed point*, the controls showed the least connectivity (to 29 regions) with those with OWL LD showing about the same (to 32 regions); and again those with dysgraphia

showed more connectivity (to 58 regions); and those with dyslexia showed the most connectivity (to 88 regions). For the *left inferior frontal seed point*, the controls had the least connectivity (to 10 regions); those with OWL LD had somewhat more (to 21 regions); those with dysgraphia had even more (to 53 regions); and those with dyslexia had the most (to 72 regions). Overall, depending on seed point of origin, those with OWL LD were either underconnected (left occipital temporal and left supramarginal) or were slightly more connected (left precuneus and left inferior gyrus) than the control group, but were always considerably less connected than were those with dysgraphia or dyslexia. Those with dysgraphia were always more connected from each seed point than were controls; and those with dyslexia were consistently overconnected compared to each of the other groups.

Pattern of connections As shown in Supplementary Table 5, the pattern of fMRI functional connectivity from each of the four seed points to which of the other brain regions showed functional connectivity was also different for each diagnostic group. That is, not only how many functional connections are made but also to where in the brain they are made may differentiate dysgraphia, dyslexia, OWL LD, and controls.

Discussion

Overview of connectome supporting word-specific spelling Compared to controls, overall those with dysgraphia made more functional connections from specific locations to other destinations and varied on the same task as to the destinations of those connections from a common location of origin. Overall, those with dyslexia showed the most functional connectivity (number of regions connected to seed point of origin) and differed most from the others in the nature of the connections (destinations). If more functional connections consume more limited resources, fewer connections may be more efficient in use of limited resources for fueling the brain's work during written word learning. Thus, the brains of those with dyslexia may be inefficient in creating functional connections during word-specific spelling judgments. In general on the same task those with OWL LD showed about the same number of functional connections as controls or fewer connections than controls depending on the seed point. Thus, those with OWL LD may not differ so much from controls on a word level spelling task not related to their hallmark syntactic-level impairment.

However, each of the SLDs showed some common functional connectivity as to destination from the seed point of origin, as well as distinctly different patterns of connectivity from each other and controls from four seed points on the same word-specific spelling task. Thus, common and unique brain bases underlie the same word-specific task for dysgraphia, dyslexia, and OWL LD and even controls. Conclusions based on this study should be restricted to developmental levels of the participants—middle childhood and early adolescence.

On the one hand, the variations in functional connectivity observed during middle childhood and early adolescence may be due to genetic variants associated with spelling (e.g., Roeske et al., 2011; Rubenstein, Matsushita, Berninger, Raskind, & Wijnsman, 2011; Schulte-Korne et al., 1998) or endophenotypes for attention

supporting language learning (e.g., Posner and Rothbart, 2007; Rubenstein et al., 2014). On the other hand, the persisting problems may be due to instruction that is not tailored to the changing nature and requirements of the curriculum during middle childhood and adolescence (e.g., see Troia, 2009), which needs to include spelling and go beyond phonology only to orthography and morphology and their interrelationships with each other. Consistent with the highly connected human connectome (Caspers et al., 2012), brain systems involved in processing word-specific spellings are not insular modules, but rather complex systems of multiple, interacting neural networks.

Conclusions

In Study 1, we first took the perspective of the educational practitioner confronted with assessment of an individual student and determining if a student has an SLD affecting written language, and if so, identifying the nature of the SLD that is relevant to planning instructional intervention. That approach confirmed unique diagnostic patterns based on impairments in levels of language and working memory components supporting language learning, but also some variation within diagnostic groups, consistent with individual differences and genetic heterogeneity in SLDs (e.g., Raskind et al., 2012), and also some commonality across groups.

Then we took the perspective of the researcher employing group statistical analyses to validate evidence-based approaches to differential diagnosis. Both distinctive and co-occurring SLDs, as predicted by the cascading levels of language framework, and common and unique impairments in working memory components, as predicted by a multi-component working memory model of language learning, were identified. Cognitive-Linguistic Translation, a higher order executive function, which works with working memory to support language learning, differentiated word-level impairment in dyslexia and syntax-level impairment in OWL LD. In sum, Study 1 provided conceptually grounded, behavioral evidence that persisting SLDs during middle childhood and adolescence can be defined, identified, and diagnosed.

Study 2 provided converging fMRI functional connectivity brain evidence that those with dysgraphia, dyslexia, OWL LD, and controls differ in neuroimaging patterns of (a) which local brain regions of origin are connected to which other brain regions, and (b) how many functional connections are made from the same local brain region to other brain regions during a common word-specific spelling task. Thus, based on behavioral and brain evidence, dysgraphia, dyslexia, and OWL LD appear to be unique SLDs, even if they share some commonalities.

Translating this research into clinical and educational practice (Mayer, 2011) will require a three-fold approach. First, translating this research to practice will require *scale up* to change policy and government regulation procedures so that a single umbrella category for learning disabilities is replaced with acknowledgement that best professional practices use evidence-based, treatment-relevant, differential diagnosis for the different kinds of SLDs. Different students may have different kinds of SLDs with different instructional needs, and within diagnostic categories there may also be some individual differences to consider in educational planning.

Second, translating research into practice requires *scale down* to grant educational professionals in local school settings professional autonomy to implement best professional practices and the rapidly expanding research knowledge about the biological bases and instructional needs for dysgraphia, dyslexia, and OWL LD flexibly and compassionately for the students with whom they work (Berninger, 2015). Legislation and government regulation are not as effective at local translation as knowledgeable professionals. Third, for this reason, translating research into practice also requires *scale out* to provide necessary professional development for all levels of a school system—from superintendents at the top to principals in local schools to teachers and all members of an interdisciplinary team in a school setting about assessing and tailoring instruction for diverse learners with and without SLDs.

Most importantly, we need to begin to ask at all levels of the system from national to state to local schools: *What Works for Whom?* Answering this question will also be pivotal to advancing scientific knowledge of how the brain learns language. Careful definitions that support sharing data across multiple sites for samples acquired in comparable ways is critical for future research not only on the SLDs investigated in the current study, but also other disorders, if research is to be translated into practice in valid and effective ways.

Acknowledgments The current study, supported by grant P50HD071764 from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) at the National Institutes of Health (NIH) to the University of Washington Learning Disabilities Research Center, has been a team effort: The first two authors contributed equally to this study. The first author recruited the sample and supervised the assessments. The second author supervised the imaging team that collected and analyzed the fMRI functional connectivity data. The last author conducted the data analyses for Study 1 and contributed to the design of cross-center Projects 1 and 3. The co-authors acknowledge the contributions of Jasmine Niedo, Roxana DelCampo, and Whitney Griffin who administered the comprehensive assessment battery, Terry Mickail who provided data base management, the imaging team involved in the larger brain imaging project, Tom Grabowski, Katie Askren, Zoe Mestre, Kevin Yagle, and Peter Boord, and Wendy Raskind who has collaborated for two decades on identifying phenotypic markers of genetic bases of written language learning disabilities. The team also gratefully acknowledges the participating children and parents' contributions to the study.

References

- Apel, K., & Apel, L. (2011). Identifying intra-individual differences in students' written language disabilities. *Topics in Language Disorders, 31*(1), 54–72.
- Baddeley, A. (2003). Working memory: Looking back. Looking forward. *Nature Reviews/Neuroscience, 4*, 829–839.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173.
- Bahr, R. H., Silliman, E. R., Berninger, V. W., & Dow, M. (2012). Linguistic pattern analysis of misspellings of typically developing writers in grades 1 to 9. *Journal of Speech, Language and Hearing Research, 55*, 1587–1599. First published on April 3, 2012 as doi:10.1044/1092-4388(2012/10-0335)
- Barnett, A., Henderson, L., Scheib, B., & Schulz, C. (2007). *Detailed assessment of speed of handwriting (DASH) copy best and fast*. London: Pearson.
- Batshaw, M., Roizen, N., & Lotrecchiano, G. (2013). *Children with disabilities* (7th ed.). Baltimore, MD: Paul H. Brookes.

- Berninger, V. (2008). Defining and differentiating dyslexia, dysgraphia, and language learning disability within a working memory model. In E. R. Silliman & M. Mody (Eds.), *Language impairment and reading disability—interactions among brain, behavior, and experience* (pp. 103–134). New York: Guilford Press.
- Berninger, V. (2009). Highlights of programmatic, interdisciplinary research on writing. *Learning Disabilities Research and Practice*, 24, 68–79.
- Berninger, V. W., & Advisory Panel. (2015). *Interdisciplinary frameworks for schools: Best professional practices for serving the needs of all students*. Washington, DC: American Psychological Association. Short Title: *Interdisciplinary Frameworks for Schools*. Companion Websites with Readings and Resources. All royalties go to Division 16 to support these websites and develop future editions.
- Berninger, V., & Abbott, R. (2013). Children with dyslexia who are and are not gifted in verbal reasoning. *Gifted Child Quarterly*, 57, 223–233. doi:10.1177/0016986213500342. Posted on PubMedCentral on 2013-09-22 15:41:46 for release August 30, 2014. NIHMSID #526583.
- Berninger, V., & Hayes, J. R. (2012). Longitudinal case studies of twenty children on writing treks in grades 1 to 5. In M. Fayol, D. Alamargot, & Berninger, V. (Eds.). *Translation of thought to written text while composing: Advancing theory, knowledge, methods, and applications* (pp. 95–179). Routledge: Psychology Press/Taylor Francis Group.
- Berninger, V., Nielsen, K., Abbott, R., Wijnsman, E., & Raskind, W. (2008a). Writing problems in developmental dyslexia: Under-recognized and under-treated. *Journal of School Psychology*, 46, 1–21.
- Berninger, V., & O'Malley May, M. (2011). Evidence-based diagnosis and treatment for specific learning disabilities involving impairments in written and/or oral language. *Journal of Learning Disabilities*, 44, 167–183.
- Berninger, V., Raskind, W., Richards, T., Abbott, R., & Stock, P. (2008b). A multidisciplinary approach to understanding developmental dyslexia within working memory architecture: Genotypes, phenotypes, brain, and instruction. *Developmental Neuropsychology*, 33, 707–744.
- Berninger, V., & Richards, T. (2010). Inter-relationships among behavioral markers, genes, brain, and treatment in dyslexia and dysgraphia. *Future Neurology*, 5, 597–617. doi:10.2217/fnl.10.22.
- Berninger, V., Winn, W., Stock, P., Abbott, R., Eschen, K., Lin, C., et al. (2008). Tier 3 specialized writing instruction for students with dyslexia. *Reading and Writing. An Interdisciplinary Journal*, 21, 95–129. Printed Springer On Line. May 15, 2007.
- Bishop, D. V. M. (2009). Specific language impairment as a language learning disability. *Child Language Teaching and Therapy*, 25, 163–165.
- Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment. *Psychological Bulletin*, 130, 858–886.
- Bowers, P., & Wolf, M. (1993). Theoretical links between naming speed, precise timing mechanisms, and orthographic skill in dyslexia. *Reading and Writing An International Journal*, 5, 69–85.
- Cahill, L., Tiberius, C., & Herring, J. (2013). PolyOrth: Orthography, phonology, and morphology in the inheritance lexicons. *Written Language and Literacy*, 16, 146–185.
- Caspers, S., Schleicher, A., Bacha-Trams, M., Palomero-Gallagher, N., Amunts, K., & Zilles, K. (2012, Feb 28). Organization of the human inferior parietal lobe based on receptor architectonics. *Cerebral Cortex*. doi:10.1093/cercor/bhs048
- Catts, H. W., Adlof, S. M., Hogan, T. P., & Ellis Weismer, S. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research*, 48, 1378–1396.
- Catts, H. W., Bridges, M. S., Little, T. D., & Tomblin, J. B. (2008). Reading achievement growth in children with language impairments. *Journal of Speech- Language and Hearing Research*, 51, 1569–1579.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. New York: MIT Press.
- Christensen, C., & Wauchope, M. (2009). Whole school literacy: Using research to create programs that build universal high levels of literate competence. In S. Rosenfield & V. Berninger (Eds.). *Implementing evidence-based interventions in school settings* (pp. 501–526). New York: Oxford University Press.
- Connelly, V., Dockrell, J., & Barnett, A. (2012). Children challenged by writing due to language and motor difficulties. In V. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 217–245). New York: Psychology Press/Taylor Francis Group.

- Crosson, B., Rao, S., Woodley, S., Rosen, A., Bobholz, J., Mayer, A., et al. (1999). Mapping of semantic, phonological, and orthographic verbal working memory in normal adults with functional magnetic resonance imaging. *Neuropsychology*, *13*, 171–187.
- Delis, D., Kaplan, E., & Kramer, J. (2001). *Delis-kaplan executive function system*. San Antonio: The Psychological Corporation/Pearson.
- Ehri, L. (1980). The role of orthographic images in learning printed words. In J. F. Kavanaugh & R. Venezky (Eds.), *Orthographic reading and dyslexia* (pp. 307–332). Baltimore, MD: University Park Press.
- Eickhoff, S. B., Heim, S., Zilles, K., & Amunts, K. (2006). Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. *NeuroImage*, *32*(2), 570–582.
- Eickhoff, S. B., Paus, T., Caspers, S., Grosbras, M. H., Evans, A., Zilles, K., & Amunts, K. (2007). Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *NeuroImage*, *36*(3), 511–521.
- Eickhoff, S., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, *25*(4), 1325–1335.
- Ellis, E. M., & Thal, D. J. (2008). Early language delay and risk for language impairment. *Perspectives on Language Learning and Education*, *15*(3), 93–100.
- Foorman, B. R., Arndt, E. J., & Crawford, E. C. (2011). Important constructs in literacy learning across disciplines. *Topics in Language Disorders*, *31*(1), 73–83.
- Greenblatt, E., Mattis, S., & Trad, P. (1990). Nature and prevalence of learning disabilities in a child psychiatric population. *Developmental Neuropsychology*, *6*, 71–83.
- Katusic, S. K., Barbaresi, W. J., Colligan, R. C., Weaver, A. L., Leibson, C. L., & Jacobsen, S. J. (2005). Case definition in epidemiologic studies of AD/HD. *Annals Epidemiology*, *15*, 430–437.
- Katusic, S. K., Colligan, R. C., Barbaresi, W. J., Schaid, D. J., & Jacobsen, S. J. (2001). Incidence of reading disability in a population-based birth cohort, 1976–1982, Rochester, Minnesota. *Mayo Clinic Proceedings*, *76*, 1081–1092.
- Katusic, S. K., Colligan, R. C., Weaver, A. L., & Barbaresi, W. J. (2009). The forgotten learning disability—Epidemiology of written language disorder in a population-based birth cohort (1976–1982), Rochester, Minnesota. *Pediatrics*, *123*, 1306–1313.
- Lefly, D., & Pennington, B. (1991). Spelling errors and reading fluency in dyslexics. *Annals of Dyslexia*, *41*, 143–162.
- Leonard, C., Eckert, M., Given, B., Berninger, V., & Eden, G. (2006). Individual differences in anatomy predict reading and oral language deficits. *Brain*, *129*, 3329–3342.
- Lovett, M., Barron, R., & Frijters, J. (2013). Word identification difficulties in children and adolescents with reading disabilities: Intervention research findings. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 329–359). New York, NY: Guilford Press.
- Lyytinen, H., Aro, M., Elklund, K., Erskine, J., Gottorm, T., Laakso, M.-L., et al. (2004). The development of children at familial risk for dyslexia: Birth to early school age. *Annals of Dyslexia*, *54*, 184–220.
- Mashburn, A. J., & Myers, S. S. (2010). Advancing research on children with speech-language impairment: An introduction to the early childhood longitudinal study—Kindergarten cohort. *Language, Speech, and Hearing Services in Schools*, *41*, 61–69.
- Mather, N., Hammill, D., Allen, E., & Roberts, R. (2004). *Test of silent word reading fluency TOSWRF*. Austin, TX: Pro-Ed.
- Mather, N., Roberts, R., Hammill, D., & Allen, E. (2008). *Test of orthographic competence (TOC)*. Austin, TX: Pro-Ed.
- Mayer, R. E. (2011). *Applying the science of learning*. Boston: Pearson.
- Möll, K., Kunze, S., Neuhoff, N., Bruder, J., & Schulte-Körne, G. (2014). Specific learning disorder: Prevalence and gender differences. *PLoS ONE*, *9*, 1–8.
- Myhill, D. (2008). Towards a linguistic model of sentence development in writing. *Language and Education*, *22*(5), 271–288.
- Nagy, W., Berninger, V., & Abbott, R. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle school students. *Journal of Educational Psychology*, *98*, 134–147.
- Nagy, W., Berninger, V., Abbott, R., Vaughan, K., & Vermeulen, K. (2003). Relationship of morphology and other language skills to literacy skills in at-risk second graders and at-risk fourth grade writers. *Journal of Educational Psychology*, *95*, 730–742.

- Nelson, N. W. (2010). *Language and literacy disorders: Infancy through Adolescence*. Boston, MA: Allyn & Bacon.
- Niedo, J., Abbott, R., & Berninger, V. (2014). Predicting levels of reading and writing achievement in typically developing, English-speaking 2nd and 5th graders. *Learning and Individual Differences*, 32C, 54–68. Published on line April 18, 2014. doi:10.1016/j.lindif.2014.03.013. NIHMS ID: NIHMS580076.
- Olson, R., Forsberg, H., Wise, B., & Rack, J. (1994). Measurement of word recognition, orthographic, and phonological skills. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities* (pp. 243–277). Brooks: Baltimore.
- Paul, R., Murray, C., Clancy, K., & Andrews, D. (1997). Reading and metaphonological outcomes in late talkers. *Journal of Speech, Language, and Hearing Research*, 40, 1037–1047.
- Pearson. (2009). *Wechsler individual achievement test* (3rd ed.). San Antonio, TX: The Psychological Corporation.
- Pennington, B. F., & Bishop, D. V. M. (2009). Relations among speech, language, and reading disorders. *Annual Review of Psychology*, 60, 283–306.
- Posner, M., & Rothbart, M. (2007). *Educating the human brain*. Washington, DC: American Psychological Association.
- Purcell, J., Turkeltaub, P. E., Eden, G. F., & Rapp, B. (2011). Examining the central and peripheral processes of written word production through meta-analysis. *Frontiers in Psychology*, 2, 1–16.
- Raskind, W., Igo, R., Chapman, N., Berninger, V., Thomson, J., Matsushita, M., et al. (2005). A genome scan in multigenerational families with dyslexia: Identification of a novel locus on chromosome 2q that contributes to phonological decoding efficiency. *Molecular Psychiatry*, 10(7), 699–711.
- Raskind, W., Peters, B., Richards, T., Eckert, M., & Berninger, V. (2012). The genetics of reading disabilities: From phenotype to candidate genes. *Frontiers in Psychology*, 3, 601. Published online 2013 January 7. doi:10.3389/fpsyg.2012.00601
- Richards, T., Aylward, E., Raskind, W., Abbott, R., Field, K., Parsons, A., et al. (2006). Converging evidence for triple word form theory in children with dyslexia. *Developmental Neuropsychology*, 30, 547–589.
- Richards, T., Berninger, V., & Fayol, M. (2009a). fMRI activation differences between 11-year-old good and poor spellers' access in working memory to temporary and long-term orthographic representations. *Journal of Neurolinguistics*, 22, 327–353.
- Richards, T., Berninger, V., Nagy, W., Parsons, A., Field, K., & Richards, A. (2005). Brain activation during language task contrasts in children with and without dyslexia: Inferring mapping processes and assessing response to spelling instruction. *Educational and Child Psychology*, 22(2), 62–80.
- Richards, T., Berninger, V., Winn, W., Swanson, H. L., Stock, P., Liang, O., et al. (2009b). Differences in fMRI activation between children with and without spelling disability on 2-back/0-back working memory contrast. *Journal of Writing Research*, 1(2), 93–123. Download the pdf from the JOWR-website.
- Roeske, D., Ludwig, K. U., Neuhoff, N., Becker, J., Bartling, J., Bruder, J., et al. (2011). First genome-wide association scan on neurophysiological 1729 endophenotypes points to trans-regulation effects on SLC2A3 in dyslexic children. *Mol 1730 Psychiatry*, 16, 97–107.
- Rubenstein, K., Matsushita, M., Berninger, V., Raskind, W., & Wijsman, E. (2011). Genome scan for spelling deficits: Effects of verbal IQ on models of transmission and trait gene localization. *Behavioral Genetics. An International Journal Devoted to the Inheritance of Behavior*, 41, 31–42. <http://www.springerlink.com/content/13017v24656mqr32/>.
- Rubenstein, K., Raskind, W., Berninger, V., Matsushita, M., Wijsman, E. (2014). Genome scan for cognitive trait loci of dyslexia: Rapid naming and rapid switching of letters, numbers, and colors. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. NIHMSID-595138.
- Saddler, B., & Graham, S. (2005). The effects of peer-assisted sentence-combining instruction on the writing performance of more and less skilled young writers. *Journal of Educational Psychology*, 97, 43–54.
- Samuelsson, S., Byrne, B., Olson, R., Hulslander, J., Wadsworth, S., Corley, R., et al. (2008). Response to early literacy instruction in the United States. *Australia, and Scandinavia: A Behavioral-Genetic Analysis, Learning and Individual Differences*, 18, 289–295.
- Scarborough, H. S. (2005). Developmental relationships between language and reading: Reconciling a beautiful hypothesis with some ugly facts. In H. W. Catts & A. G. Kamhi (Eds.), *The connections between language and reading abilities* (pp. 3–24). Mahwah, NJ: Erlbaum.

- Schneider, W., & Shiffrin, R. (1977). Controlled and automatic human information processing: Detection, search, and attention. *Psychological Review*, *84*, 1–66.
- Schulte-Korne, G., Grimm, T., Nothen, M. M., Muller-Myhsok, B., Cichon, S., Vogt, I. R., et al. (1998). Evidence for linkage of spelling disability to chromosome 15. *American Journal of Human Genetics*, *63*, 279–282.
- Scott, C. M. (2010). Assessing expository texts produced by school-age children and adolescents. In M. A. Nippold & C. M. Scott (Eds.), *Expository discourse in children, adolescents, and adults* (pp. 191–214). New York: Psychology Press.
- Scott, C. M. (2011). Assessment of language and literacy: A process of hypothesis testing. *Topics in Language Disorders*, *31*(1), 24–39.
- Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluations of language fundamentals 4th edition: Examiner's manual*. San Antonio, TX: Harcourt Assessment Inc.
- Shriffin, R., & Scheider, W. (1977). Controlled and automatic processing II: Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 70–120.
- Silliman, E., & Berninger, V. (2011). Cross-disciplinary dialogue about the nature of oral and written language problems in the context of developmental, academic, and phenotypic profiles. *Topics in Language Disorders*, *31*, 6–23. Free access at http://journals.lww.com/topicsinlanguagedisorders/Fulltext/2011/01000/Cross_Disciplinary_Dialogue_about_the_Nature_of.3.aspx.
- Silliman, E. R., & Mody, M. (2008). Individual differences in oral language and reading: It's a matter of individual differences. In M. Mody & E. R. Silliman (Eds.), *Brain, behavior, and learning in language and reading disorders* (pp. 349–386). New York: Guilford Press.
- Silliman, E. R., & Scott, C. M. (2009). Research-based oral language intervention routes to the academic language of literacy: Finding the right road. In S. Skibbe, L. E., Grimm, K. J., Stanton-Chapman, T. L., Justice, L. M., Pence, K. L., & Bowles, R. P. (2008). Reading trajectories of children with language difficulties from preschool through fifth grade. *Language, Speech, and Hearing Services in Schools*, *39*, 475–486.
- St. Sauver, J. L., Katusic, S. K., Barbaresi, W. J., Colligan, R. C., & Jacobsen, S. J. (2001). Boy/girl differences in risk for reading disability: Potential clues? *American Journal of Epidemiology*, *154*, 787–794.
- Stoekel, R. E., Colligan, R. C., Barbaresi, W. J., Weaver, A. L., Killian, J. M., & Katusic, S. K. (2013). Early speech-language impairment and risk for written language disorder: A population-based study. *Journal Developmental Behavior Pediatrics*, *34*, 38–44.
- Thal, D. J., Bates, E., Goodman, J., & Jahn-Samilo, J. (1997). Continuity of language abilities: An exploratory study of late-and early-talking toddlers. *Developmental Neuropsychology*, *13*, 239–273.
- Thal, D. J., & Katich, J. (1996). Predicaments in early identification of specific language impairment: Does the early bird catch the worm? In K. N. Cole, P. S. Dale, & D. J. Thal (Eds.), *Assessment of communication and language* (pp. 1–28). Baltimore: Paul H. Brookes.
- Toga, A., Thompson, P., Mori, S., Amunts, K., & Zilles, K. (2006). Towards multimodal atlases of the human brain. *Nature Reviews Neuroscience*, *7*, 952–966.
- Torgesen, J., Wagner, R., & Rashotte, C. (1999). *Test of word reading efficiency*. Austin, TX: Pro-Ed.
- Treiman, R., & Kessler, B. (2014). *How children learn to write words*. New York: Oxford University Press.
- Troia, G. (Ed.). (2009). *Instruction and assessment for struggling writers: Evidence-based practices*. New York: Guilford.
- van Viersen, S., Kroesbergen, E., Slot, E., & de Bree, E. (2014). High reading skills mask dyslexia in gifted children. *Journal of Learning Disabilities*. Epub ahead of print.
- Vellutino, F., Scanlon, D., & Tanzman, M. (1991). Bridging the gap between cognitive and neuropsychological conceptualizations of reading disabilities. *Learning and Individual Differences*, *3*, 181–203.
- Venezky, R. (1970). *The structure of English orthography*. The Hague: Mouton.
- Venezky, R. (1999). *The American way of spelling*. New York: Guilford.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *The comprehensive test of phonological processing*. Austin, TX: Pro-Ed.
- Wechsler, D. (2003). *Wechsler intelligence scale for children, 4th edition (WISC-IV)*. San Antonio, TX: The Psychological Corporation.
- Wolf, M., & Denckla, M. (2005). *RAN/RAS rapid automatized naming and rapid alternating stimulus tests*. Austin, TX: Pro-Ed.

- Woodcock, R., McGrew, K., & Mather, N. (2001a). *Woodcock-Johnson III psychoeducational cognitive test battery*. Itasca, IL: Riverside.
- Woodcock, R., McGrew, K., & Mather, N. (2001b). *Woodcock-Johnson III achievement battery*. Itasca, IL: Riverside.
- Yoshimasu, K., Barbaresi, W. J., Colligan, R., Killian, J., Voigt, R. G., Weaver, A., et al. (2011). Written language disorder in children with and without attention-deficit/hyperactivity disorder in a population-based birth cohort. *Pediatrics*, *128*(3), e605-12. *Pediatrics* 2010; *126*, e788–e795.
- Yoshimasu, K., Barbaresi, W. J., Colligan, R. C., Killian, J. M., Voigt, R. G., Weaver, A. L., et al. (2012). Gender, ADHD, and reading disability in a population-based birth cohort. *Pediatrics*, *126*, e788–e795.